

301 TTAAACGATGGAATCAGAAAGCATTCGAAAGCTAGGGTTTCGAAACCGCCGCAATTCGGT 360
301 TTAAACGATGGAATCAGAAAGCATTCGAAAGCTAGGGTTTCGAAACCGCCGCAATTCGGT 360
361 TTCAAGGACGAGCGCTTTAAATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
361 TTCAAGGACGAGCGCTTTAAATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
421 TCTAAACCTCGGCTTGAAGAGAGTACAAATGAAGGCTCTTGATGATGAAGAAAGCTACA 480
421 TCTAAACCTCGGCTTGAAGAGAGTACAAATGAAGGCTCTTGATGATGAAGAAAGCTACA 480
481 GTCAATCACTGATGTTCTTGGGTAAGGTTCTGCGGCTCTCTGTGTAATTCAGAAAGGT 540
481 GTCAATCACTGATGTTCTTGGGTAAGGTTCTGCGGCTCTCTGTGTAATTCAGAAAGGT 540
541 GGTGAGACTGAGATGATCTTCTGCGGTTGAGAGGCTCTGCTTAAGAGAGGTTGCCTAAG 600
541 GGTGAGACTGAGATGATCTTCTGCGGTTGAGAGGCTCTGCTTAAGAGAGGTTGCCTAAG 600
601 TCGTTTAACGAAAGATGATGTTTAAATGAGGCTGCGGTTCTGATGATGATGATGATGATGAT 660
601 TCGTTTAACGAAAGATGATGTTTAAATGAGGCTGCGGTTCTGATGATGATGATGATGATGAT 660
661 GCTATGCAATGATGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 GCTATGCAATGATGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 AACCTTTTACAG 780
721 AACCTTTTACAG 780
781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTGAGAGCTACTGCTTACCGCTTGT 840
781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTGAGAGCTACTGCTTACCGCTTGT 840
841 GATGATTAACGCTCGAAGAACTAAATGTTTAAAGCGTGTGCGAATAATTTTGTGCT 900
841 GATGATTAACGCTCGAAGAACTAAATGTTTAAAGCGTGTGCGAATAATTTTGTGCT 900
901 GTTGAAGAGAGTGAAGATCAGCTCTTGTGAGGAGTTTGAACCGGTGAAGAGTTTATGAT 960
901 GTTGAAGAGAGTGAAGATCAGCTCTTGTGAGGAGTTTGAACCGGTGAAGAGTTTATGAT 960
961 GAGCGCTTTTAAAGATGAGAGCTGCTGAGAGAGTTGATCTTTTGAAGTACCCCAAGC 1020
961 GAGCGCTTTTAAAGATGAGAGCTGCTGAGAGAGTTGATCTTTTGAAGTACCCCAAGC 1020
1021 AATATTCAGCAGAGTCAATTTGAAGTTTAAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 1080
1021 AATATTCAGCAGAGTCAATTTGAAGTTTAAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 1080
1081 TTTATTTGTAAGAAAGCACAACCTTTTAAAGAGTCTGATTAAGCAATTCAGCAACTTCAG 1140
1081 TTTATTTGTAAGAAAGCACAACCTTTTAAAGAGTCTGATTAAGCAATTCAGCAACTTCAG 1140
1141 CAGCGCTTAAGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCT 1200
1141 CAGCGCTTAAGATGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCT 1200
1201 GAGATGAGCTTGGTCTGAGAAAGGAGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 1260
1201 GAGATGAGCTTGGTCTGAGAAAGGAGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 1260
1261 TGCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
1261 TGCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
1321 GAGTTTGTGTTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
1321 GAGTTTGTGTTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
1381 TTGTTGAAACCTGTTGAGAGGAGTGTCTTTCTTCTGAGTTTCAAGAGACCAAAAGATTA 1440
1381 TTGTTGAAACCTGTTGAGAGGAGTGTCTTTCTTCTGAGTTTCAAGAGACCAAAAGATTA 1440

1381 TTGTTGAAACCTGTTGAGAGGAGTGTCTTTCTTCTGAGTTTCAAGAGACCAAAAGATTA 1440
1441 AATATTAACCTGAGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
1441 AATATTAACCTGAGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
1501 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
1501 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
1561 CATGTAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 CATGTAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 AGAACTGAGGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 1680
1621 AGAACTGAGGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 1680
1681 AACATGATGAGGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCA 1740
1681 AACATGATGAGGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCA 1740
1741 GAAACTTTGAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1741 GAAACTTTGAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1801 GAAACTTTGAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1801 GAAACTTTGAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1861 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
1861 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
1921 AGCTCATCTTTTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 1980
1921 AGCTCATCTTTTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 1980
1981 GGTGAGTCAAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCA 2040
1981 GGTGAGTCAAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCA 2040
2041 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2041 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2101 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
2101 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
2161 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
2161 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
2221 GTGACAGTCTCAGCAGATGAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGT 2280
2221 GTGACAGTCTCAGCAGATGAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGT 2280
2281 TGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
2281 TGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
2341 AGATGAGAGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2341 AGATGAGAGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2401 TCATTA 2406
2401 TCATTA 2406

RESULT 2


```

Db      1801 GAAACACTGTTGAAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Qy      1861 GCTGCTGCTGTGGCAATTGGACTGATTTCACTGTTCAGCCAGAAAGTATTTTCTTAAAGC 1920
Db      1861 GCTGCTGCTGTGGCAATTGGACTGATTTCACTGTTCAGCCAGAAAGTATTTTCTTAAAGC 1920
Qy      1921 AGCTCACTCTTTTCAAGCAAGATATGATGTTCTTCTATGAAATCTGATGTGCTACATA 1980
Db      1921 AGCTCACTCTTTTCAAGCAAGATATGATGTTCTTCTATGAAATCTGATGTGCTACATA 1980
Qy      1981 GGGTCAGTCAGACCTGACATTCAGAAAGCACTTCCAGAAATGAAATGCTAGACTGACAG 2040
Db      1981 GGGTCAGTCAGACCTGACATTCAGAAAGCACTTCCAGAAATGAAATGCTAGACTGACAG 2040
Qy      2041 AATATATGATCCAAATGCGCAAGAAATTAAGTCTCTGGCTTTTGGGCTCTGATCACCGATA 2100
Db      2041 AATATATGATCCAAATGCGCAAGAAATTAAGTCTCTGGCTTTTGGGCTCTGATCACCGATA 2100
Qy      2101 GAAATGTTACCAAGGTTTGGATGAGGCGAAATGCTGAAGATTTGGAAGTGAAGTGAAGCAGCT 2160
Db      2101 GAAATGTTACCAAGGTTTGGATGAGGCGAAATGCTGAAGATTTGGAAGTGAAGTGAAGCAGCT 2160
Qy      2161 GAAACTGCGCAGCTTGGGTTGTTATGATTAATACACTGTGAAACTATCTGTGACAT 2220
Db      2161 GAAACTGCGCAGCTTGGGTTGTTATGATTAATACACTGTGAAACTATCTGTGACAT 2220
Qy      2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGTGTGAAGCAACTGTGAAGAGTCTGCT 2280
Db      2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGTGTGAAGCAACTGTGAAGAGTCTGCT 2280
Qy      2281 TGTCTATCTGATTTGGTTGTTCAACGAAACAAATGCTACTGATGTGAGAACTTACACAA 2340
Db      2281 TGTCTATCTGATTTGGTTGTTCAACGAAACAAATGCTACTGATGTGAGAACTTACACAA 2340
Qy      2341 AGATAGCAAGTTTCTGTGTCCAAGTCAGGCTGGAATACTGAAAGCTCTGTCTTGA 2400
Db      2341 AGATAGCAAGTTTCTGTGTCCAAGTCAGGCTGGAATACTGAAAGCTCTGTCTTGA 2400
Qy      2401 TCATTA 2406
Db      2401 TCATTA 2406

RESULT 3
US-10-600-070-128
/ Sequence 128, Application US/10600070
/ Publication No. US20040139500A1
/ GENERAL INFORMATION:
/ APPLICANT: Osteoryoung, Katherine W.
/ APPLICANT: Vitsha, Stanislav
/ APPLICANT: Koshmarova, Olga A.
/ APPLICANT: Gao, Hong
/ TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
/ FILE REFERENCE: Use
/ CURRENT APPLICATION NUMBER: US/10/600,070
/ NUMBER OF SEQ ID NOS: 206
/ SOFTWARE: Patent version 3.2
/ SEQ ID NO 128
/ LENGTH: 2406
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ US-10-600-070-128

Query Match      99.7%; Score 2399.6; DB 7; Length 2406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      61 CGGCGACGACAAAGCTCCGAGTACCAACACCTTACAACTATCTGCTCCGACG 120
Db      61 CGGCGACGACAAAGCTCCGAGTACCAACACCTTACAACTATCTGCTCCGACG 120
Qy      121 AATGGGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 180
Db      121 AATGGGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 180
Qy      181 TTGCGACGCGACCAACCGCACTCGTCTCTGCGACCACTATGATGCTGCC 240
Db      181 TTGCGACGCGACCAACCGCACTCGTCTCTGCGACCACTATGATGCTGCC 240
Qy      241 GAAAGCAGTCCCGCATTCCTCAATTTCTACAGATTTAGAGCTCAACACATTT 300
Db      241 GAAAGCAGTCCCGCATTCCTCAATTTCTACAGATTTAGAGCTCAACACATTT 300
Qy      301 TTAAACGATGAAATCAAGAGCATTCGAAGCTTTCGAAACCGCGCAATTCGGT 360
Db      301 TTAAACGATGAAATCAAGAGCATTCGAAGCTTTCGAAACCGCGCAATTCGGT 360
Qy      361 TTCAGGAGAGCGCTTAAACGCGGAGACAGATTTCTCAAGCTCTTGCAGAACTTG 420
Db      361 TTCAGGAGAGCGCTTAAACGCGGAGACAGATTTCTCAAGCTCTTGCAGAACTTG 420
Qy      421 TCTAATCTCTCGCTAGAAAGAGTAAATGAAGTCTTCTGATGATGAAGAGCTACA 480
Db      421 TCTAATCTCTCGCTAGAAAGAGTAAATGAAGTCTTCTGATGATGAAGAGCTACA 480
Qy      481 GTCATCACTGATTTCTTGGATTAAGTTCCTGGGCTCTCTGTATTTGCAAGAGT 540
Db      481 GTCATCACTGATTTCTTGGATTAAGTTCCTGGGCTCTCTGTATTTGCAAGAGT 540
Qy      541 GGTGAGACTGAGATTAATCTGCGGTTGGTGAAGCTCTGCTTAAGAGAGTTCCTAAG 600
Db      541 GGTGAGACTGAGATTAATCTGCGGTTGGTGAAGCTCTGCTTAAGAGAGTTCCTAAG 600
Qy      601 TCGTTTAAACAAGATGCTTTAGTTATGCGCTTTCGATGCTGAGAGAT 660
Db      601 TCGTTTAAACAAGATGCTTTAGTTATGCGCTTTCGATGCTGAGAGAT 660
Qy      661 GCTATGCGATTTGATCACTGATTTATTAAGTGTGAGAGTTCGAGAGTTCG 720
Db      661 GCTATGCGATTTGATCACTGATTTATTAAGTGTGAGAGTTCGAGAGTTCG 720
Qy      721 AAGCTTTTACAGAGAGAGAGCAATAGCTTGAACCGGATTTAGTGAACAATGAT 780
Db      721 AAGCTTTTACAGAGAGAGAGCAATAGCTTGAACCGGATTTAGTGAACAATGAT 780
Qy      781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTGTGAGCTACTTGGCTTGGT 840
Db      781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTGTGAGCTACTTGGCTTGGT 840
Qy      841 GATGATTAACGCTCGGAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTGTGCT 900
Db      841 GATGATTAACGCTCGGAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTGTGCT 900
Qy      901 GTTGAAGAGGTGAGACATCAAGCTCTGTTGGGGGTTTACCGGTAGAAATTAAGAT 960
Db      901 GTTGAAGAGGTGAGACATCAAGCTCTGTTGGGGGTTTACCGGTAGAAATTAAGAT 960
Qy      961 GAGCGCTTTTACGAATGACAGCTGAGCAGTGTGATCTTTTGTAGTACCCAGC 1020
Db      961 GAGCGCTTTTACGAATGACAGCTGAGCAGTGTGATCTTTTGTAGTACCCAGC 1020
Qy      1021 AATATTCAGACAGATCATTTGAAGTTTACGAAGTTTGAAGCTCTTGTGGCTCAAGCT 1080
Db      1021 AATATTCAGACAGATCATTTGAAGTTTACGAAGTTTGAAGCTCTTGTGGCTCAAGCT 1080
Qy      1081 TTTATTTGTAAGAGCAACCTTTTACAGAGCTGATTAAGAAATTCAGCAACTTCAG 1140
Db      1081 TTTATTTGTAAGAGCAACCTTTTACAGAGCTGATTAAGAAATTCAGCAACTTCAG 1140

```


QY 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACGGAATAATATGG 1200
 DB 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACGGAATAATATGG 1200
 QY 1201 GAGATGACTTCGGCTAGAGAAAGGGAATCTGTGCACTGCTTAATAGGCAAAATTGATGAA 1260
 DB 1201 GAGATGACTTCGGCTAGAGAAAGGGAATCTGTGCACTGCTTAATAGGCAAAATTGATGAA 1260
 QY 1261 TGGCGTATGCTGCTGCTTGAACAGTGAAGATTCACAAATATAGAAATCCAGCTATATGG 1320
 DB 1261 TGGCGTATGCTGCTGCTTGAACAGTGAAGATTCACAAATATAGAAATCCAGCTATATGG 1320
 QY 1321 GAGTTGTTTGGAGAAATCAATCGATGATCAATGATGATCTCCCTGCACTATGCAAA 1380
 DB 1321 GAGTTGTTTGGAGAAATCAATCGATGATCAATGATGATCTCCCTGCACTATGCAAA 1380
 QY 1381 TTGTTGAAAACCTGTTGGCAGGGGTTGCTTTCTAGTTTCAAGAGACACCAAAAGATPAA 1440
 DB 1381 TTGTTGAAAACCTGTTGGCAGGGGTTGCTTTCTAGTTTCAAGAGACACCAAAAGATPAA 1440
 QY 1441 AAATTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTAATCTTGAAAGAGTG 1500
 DB 1441 AAATTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTAATCTTGAAAGAGTG 1500
 QY 1501 GAGTAGTTGAGGGTCTCCTTTAGCTGCTGCTGCACTATGCAAGATTTGAGCCGAG 1560
 DB 1501 GAGTAGTTGAGGGTCTCCTTTAGCTGCTGCTGCACTATGCAAGATTTGAGCCGAG 1560
 QY 1561 CATGTGAAGCTAGTGTCTATGCAAGCACTGCAAGAAATTTTCCCTTCGCTATACAGAT 1620
 DB 1561 CATGTGAAGCTAGTGTCTATGCAAGCACTGCAAGAAATTTTCCCTTCGCTATACAGAT 1620
 QY 1621 AGAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
 DB 1621 AGAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
 QY 1681 AACAAATGAGCGCGTATGAGCTGAGCTGCTTTATGCAAGAAAGCTGTAAGACCTCT 1740
 DB 1681 AACAAATGAGCGCGTATGAGCTGAGCTGCTTTATGCAAGAAAGCTGTAAGACCTCT 1740
 QY 1741 GAAAACCTTTGAAAATGATGATGCAATTCGAGCTGCGGCTCTCAGAGATGAGCTTAT 1800
 DB 1741 GAAAACCTTTGAAAATGATGATGCAATTCGAGCTGCGGCTCTCAGAGATGAGCTTAT 1800
 QY 1801 GAAACTACTGTTGAAATGTCCTGCTGATGATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
 DB 1801 GAAACTACTGTTGAAATGTCCTGCTGATGATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
 QY 1861 GCTGCTGCTGCTGCAATGCACTGATTCCTGCTGAGCGAAGATTTTCTTAAAGC 1920
 DB 1861 GCTGCTGCTGCTGCAATGCACTGATTCCTGCTGAGCGAAGATTTTCTTAAAGC 1920
 QY 1921 AGCTCAATCTTTTCAACGCAAGATATGTTTCTTCTATGAAATCTGATGCTGATACATA 1980
 DB 1921 AGCTCAATCTTTTCAACGCAAGATATGTTTCTTCTATGAAATCTGATGCTGATACATA 1980
 QY 1981 GGGTCACTGAGCTGCAATTCGAAAGCACTTCCGAAATGAGTGTGAGACTGCAAG 2040
 DB 1981 GGGTCACTGAGCTGCAATTCGAAAGCACTTCCGAAATGAGTGTGAGACTGCAAG 2040
 QY 2041 AAATATGATTCGAAGTGGCAAGATTAAGTCTGCTGCTTTGGGCTGATACCGGATA 2100
 DB 2041 AAATATGATTCGAAGTGGCAAGATTAAGTCTGCTGCTTTGGGCTGATACCGGATA 2100
 QY 2101 GAAATGTTTACAGAGTTTGAATGCGCAATGCTGAAGATTTGCACTGCAAGAGCACT 2160
 DB 2101 GAAATGTTTACAGAGTTTGAATGCGCAATGCTGAAGATTTGCACTGCAAGAGCACT 2160
 QY 2161 GAAACTGCGAGCTTGGGTTGTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220
 DB 2161 GAAACTGCGAGCTTGGGTTGTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220
 QY 2221 GTGACAGTCTCAGCAAGTGAACCCGCTGCTGCTGAGAAACAACCTGCGAGAGTCTGCT 2280

DB 2221 GTGACAGTCTCAGCAAGTGAACCCGCTGCTGCTGAGAACCACTGAGAGTCTGCT 2280
 QY 2281 TGTCTATCTGATTTGGTCTTATCCGAAAGAAACAAATGCTACTGATGTCAAGACTTACACACA 2340
 DB 2281 TGTCTATCTGATTTGGTCTTATCCGAAAGAAACAAATGCTACTGATGTCAAGACTTACACACA 2340
 QY 2341 AGATPAGAAATTTTCTGCTGCTGCAAGTCAAGGTTGGAATCACTGAAAGCTGTTCTTGA 2400
 DB 2341 AGATPAGAAATTTTCTGCTGCTGCAAGTCAAGGTTGGAATCACTGAAAGCTGTTCTTGA 2400
 QY 2401 TCATPAA 2406
 DB 2401 TCATPAA 2406
 RESULT 4
 US-10-600-070-130
 ; Sequence 130, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OsteYoung, Katherine W.
 ; APPLICANT: Valthe, Stanislaw
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hongbo
 ; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20
 ; NUMBER OF SEQ. ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ. ID NO 130
 ; LENGTH: 2637
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-600-070-130
 Query Match 99.7%; Score 2399.6; DB 7; Length 2637;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGGAAGCTGAGTCAAGCTGCGGCAATGCTGCTCCCATTCGAATTATGCGGATTCACA 60
 DB 114 ATGGAAGCTGAGTCAAGCTGCGGCAATGCTGCTCCCATTCGAATTATGCGGATTCACA 173
 QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCAACACCTCTACAACTATGCTCCGCGAGC 120
 DB 174 CCGGCGACGACAAAGCTCCGACGTAAGCAACACCTCTACAACTATGCTCCGCGAGC 233
 QY 121 AAATGAGCGGACCGTCTTCTCTCGACTTCATTTCACTCGGATTCCTCTCTCTCC 180
 DB 234 AAATGAGCGGACCGTCTTCTCTCGACTTCATTTCACTCGGATTCCTCTCTCTCC 293
 QY 181 TTGCGCAACCGGCAACCAACCGGCACTGCTGCTCTGCAACATCTATGATGCTGCC 240
 DB 294 TTGCGCAACCGGCAACCAACCGGCACTGCTGCTCTGCAACATCTATGATGCTGCC 353
 QY 241 GAAAGCAGCTCCCATTCCTCATTTCTACAGATATAGAGCTCAACACATTTTC 300
 DB 354 GAAAGCAGCTCCCATTCCTCATTTCTACAGATATAGAGCTCAACACATTTTC 413
 QY 301 TTAAACGATGAAATCAAGAGCAATTCGAAGCTTGAAGCTTGAACCGCGCAATTCGGT 360
 DB 414 TTAAACGATGAAATCAAGAGCAATTCGAAGCTTGAAGCTTGAACCGCGCAATTCGGT 473
 QY 361 TTCAAGGAGAGAGCTTTATCAAGCGGAGAGAGATTCCTCAAGCTCTTGGCAACTCG 420
 DB 474 TTCAAGGAGAGAGCTTTATCAAGCGGAGAGAGATTCCTCAAGCTCTTGGCAACTCG 533
 QY 421 TCTAATCTCTGCTCTAAGAGAGTCAATGAAAGTCTTCTGATGATGAAGAGCTACA 480
 DB 534 TCTAATCTCTGCTCTAAGAGAGTCAATGAAAGTCTTCTGATGATGAAGAGCTACA 593

Db 2095 GGCTACGTACAGAGCTGACGATTCAGAACCACTTCCAGAAATGATGCTAGAGCTGACAG 2154
 QY 2041 AATATATGATTCAGAAAGTGGCAGAAAGATTAACTCTGGCTTTGGGCTGATCACCGCAT 2100
 Db 2155 AATATGATTCAGAAAGTGGCAGAAAGATTAACTCTGGCTTTGGGCTGATCACCGCAT 2214
 QY 2101 GAAATGTTACAGAGGTTTGGATGAGCGAATGCTGAAAGATTGAACTGACAGAGCACT 2160
 Db 2215 GAAATGTTACAGAGGTTTGGATGAGCGAATGCTGAAAGATTGAACTGACAGAGCACT 2274
 QY 2161 GAAACTGCGCAGCTGGGTTGGTTTATGATTAATACATGTTGAAACTATCTGTTGACACT 2220
 Db 2275 GAAACTGCGCAGCTGGGTTGGTTTATGATTAATACATGTTGAAACTATCTGTTGACACT 2234
 QY 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGGTGAACCACTGAGAGAGTCTGCT 2280
 Db 2335 GTGACAGTCTCAGCAGATGGAACCCGCTCTGGTGAACCACTGAGAGAGTCTGCT 2294
 QY 2281 TGTCTATCTGATTTGGTTTCATCCAGAAACAAATGCTATGATGTCAGAACCTTACACAA 2340
 Db 2395 TGTCTATCTGATTTGGTTTCATCCAGAAACAAATGCTATGATGTCAGAACCTTACACAA 2454
 QY 2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAAATCACTGAAGGCTCTGTTCTTGA 2400
 Db 2455 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAAATCACTGAAGGCTCTGTTCTTGA 2514
 QY 2401 TCATTA 2406
 Db 2515 TCATTA 2520

RESULT 6

US-10-600-070-3
 ; Sequence 3, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitsha, Stanislav
 ; APPLICANT: Kosharova, Olga A.
 ; APPLICANT: Geo, Hong
 ; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 3
 ; LENGTH: 3667
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-600-070-3

Query Match 71.7%; Score 1724; DB 7; Length 3667;

Best Local Similarity 79.2%; Pred. No. 0;

Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

QY 1 ATGGAAGCTTGTAGTCACTGCGGATTTGCTCTCCCATTCATTAATGCGATTACCA 60
 Db 481 ATGGAAGCTTGTAGTCACTGCGGATTTGCTCTCCCATTCATTAATGCGATTACCA 540
 QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCACAACACCTCTCACTATCTGTCGCGAGC 120
 Db 541 CCGGCGACGACAAAGCTCCGACGTAAGCACAACACCTCTCACTATCTGTCGCGAGC 600
 QY 121 AAATGGCGCAGCGCTTCTCTCGACTTCATTAATTCACCTCCGATTCCTCTCTCC 180
 Db 601 AAATGGCGCAGCGCTTCTCTCGACTTCATTAATTCACCTCCGATTCCTCTCTCC 660
 QY 181 TTGGCCACCGCACAACACCGACACTCTGCTCTCTGCGACCATTAATGATGCTCC 240
 Db 661 TTGGCCACCGCACAACACCGACACTCTGCTCTCTGCGACCATTAATGATGCTCC 720

QY 241 GAAAGCCAGCTGCCCATCCCATTTGATTTCTACAGATTTAGAGCTCAACACACTTC 300
 Db 721 GAAAGCCAGCTGCCCATCCCATTTGATTTCTACAGATTTAGAGCTCAACACACTTC 780
 QY 301 TTAACCGATGGAATCGAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCCGT 360
 Db 781 TTAACCGATGGAATCGAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCCGT 840
 QY 361 TTCAGCGACGACGCTTTAATTCAGCCGAGACAGATTCCTCAAGCTCTTCCGAACTCTG 420
 Db 841 TTCAGCGACGACGCTTTAATTCAGCCGAGACAGATTCCTCAAGCTCTTCCGAACTCTG 900
 QY 421 TCTAATCTCGGCTCTAGAAAGATCAATGAAGTCTTCTTGATGATGAAGAGCTACA 480
 Db 901 TCTAATCTCGGCTCTAGAAAGATCAATGAAGTCTTCTTGATGATGAAGAGCTACA 960
 QY 481 GTCATCACTGATGCTTCCCTGGAT----- 504
 Db 961 GTCATCACTGATGCTTCCCTGGATTAAGTAAAGTAAATTCGATTCGGAATTAAGTTCTTC 1020
 QY 505 -----AGGTTCCCTGGAG 518
 Db 1021 GTTTTAATTCATGAATGGAATTAAGAAAGAACTTTAATCTAGTGAAGGTTCTGAGG 1080
 QY 519 TCTCTGTATTCAGAAAGTGTGAGACTGAGATAGTTCTTCCGCTTGTGAGGCTCT 578
 Db 1081 TCTCTGTATTCAGAAAGTGTGAGACTGAGATAGTTCTTCCGCTTGTGAGGCTCT 1140
 QY 579 GCTTAAGAGAGGTTCCCTAAGTCTTAAAGCAAGATGTTTAAAGTATGAGCGCTGC 638
 Db 1141 GCTTAAGAGAGGTTCCCTAAGTCTTAAAGCAAGATGTTTAAAGTATGAGCGCTGC 1200
 QY 639 GTTTCGATGCTCCAGGAGATGCTATGAGCATTCACCTGATTTAATTAATCTGTTA 698
 Db 1201 GTTTCGATGCTCCAGGAGATGCTATGAGCATTCACCTGATTTAATTAATCTGTTA 1260
 QY 699 TGAATTTGTTAGAAAGCTTTGAAGCTTTAC----- 730
 Db 1261 TGAATTTGTTAGAAAGCTTTGAAGCTTTACAGTATGTTGACTGCTTGTGTAATTC 1320
 QY 731 -----A 731
 Db 1321 ACGAGGTTGCTTTAATGAACCTTCTGATTTGATTTGATTTGATTTGATTTGATTTG 1380
 QY 732 GAGAGAAAGAGCAAGTACCTTGCAACGATTTACGTCACAATTTGATGAGACTTTGA 791
 Db 1381 GAGAGAAAGAGCAAGTACCTTGCAACGATTTACGTCACAATTTGATGAGACTTTGA 1440
 QY 792 AGAGATCACTCCGCTTATGCTTTGAGACTTACGCTTACGCTTGTGATGATTAAGC 851
 Db 1441 AGAGATCACTCCGCTTATGCTTTGAGACTTACGCTTACGCTTGTGATGATTAAGC 1500
 QY 852 TCGGAAAGACTAATAGTTTAAGCGGTGCGGATATTTTGTGCTTTGAGAGAG 911
 Db 1501 TCGGAAAGACTAATAGTTTAAGCGGTGCGGATATTTTGTGCTTTGAGAGAG 1560
 QY 912 TGGAGATCAGCTCTTGTGGGGTTTGAACCGGTGAGAGTTATGAATGAGGCGTTTT 971
 Db 1561 TGGAGATCAGCTCTTGTGGGGTTTGAACCGGTGAGAGTTATGAATGAGGCGTTTT 1620
 QY 972 ACGAATGACAGCTGTAGC----- 991
 Db 1621 ACGAATGACAGCTGTAGCATTAATTAATTTTAAATTTCTTAAAGCA 1680
 QY 992 -----AGTTGATCTTTTGTGA 1008
 Db 1681 TGATATTAATTAAGTTTCTCATTTTAATGATGTTGTGATGATGATGATGATGATG 1740
 QY 1009 GCTAACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGACCTTGCTTT 1068
 Db 1741 GCTAACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGACCTTGCTTT 1800
 QY 1069 GTGGCTCAAGCTTTTATGTTGTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTC 1128

```

Db      1801 GTGGCTCAAGCTTTATTTAGTAAGAACCAACCTTTTACAGAGATGCTGATTAAGCAATTC
Qy      1129 CAGCAATTCAGAGAGCTAAGTAATGCTATGAGAAATTCCTGAGATGCTATGATAGACA
Db      1861 CAGCAATTCAGAGAGCTAAGTAATGCTATGAGAAATTCCTGAGATGCTATGATAGACA
Qy      1189 CGGAATTAATGGAGATAGACTTGGCTGATGAAAGGGAATCTGTGACCTGCTTAAGGC
Db      1921 CGGAATTAATGGAGATAGACTTGGCTGATGAAAGGGAATCTGTGACCTGCTTAAGGC
Qy      1249 AAGTGTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT
Db      1981 AAGTGTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT
Qy      1309 CAGCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT
Db      2041 CAGCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT
Qy      1369 GAGCTATGCAAAATTTGTTGAAACCTGGTTGGCAGAGGGTGTCTTCTAGGTTCAAGAC
Db      2101 GAGCTATGCAAAATTTGTTGAAACCTGGTTGGCAGAGGGTGTCTTCTAGGTTCAAGAC
Qy      1429 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGAT
Db      2161 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGAT
Qy      1489 TTGGAAGAAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      2221 TTGGAAGAAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1549 ATTGAGCCGAGAGATGTAAGAGCTAGTATGCAAGGCACTGCAAGAAATTTTCTCTCC
Db      2281 ATTGAGCCGAGAGATGTAAGAGCTAGTATGCAAGGCACTGCAAGAAATTTTCTCTCC
Qy      1609 CGCTATACAGATGAAACTCGGCTGAAACCAAGGATGCAAGAGACAGTGTATGATGTA
Db      2341 CGCTATACAGATGAAACTCGGCTGAAACCAAGGATGCAAGAGACAGTGTATGATGTA
Qy      1669 GATCTGTTGGTAACTATGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      2401 GATCTGTTGGTAACTATGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1729 GTAAGACCCCTGTAATACTTGAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAG
Db      2461 GTAAGACCCCTGTAATACTTGAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAG
Qy      1789 AGTAGCGTTGATGAACTACTGTTGAATGCTCGTGTGATATGTTAAAGAGGCAAGT
Db      2521 AGTAGCGTTGATGAACTACTGTTGAATGCTCGTGTGATATGTTAAAGAGGCAAGT
Qy      1849 GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Db      2581 GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Qy      1909 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATGATGATGATGATGAT
Db      2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATGATGATGATGAT
Qy      1969 GTGCTACCA-----1978
Db      2701 GTGCTACCAATAGTATGATTAATGATGCAATTTTCATATCTGATGCTCAAAATA
Qy      1979 -----1978
Db      2761 TGCTGTTTGTGAGCTAAGACATAGTCCCACTTATACATGTCCTCAAAAGTGTAGC
Qy      1979 -----1978
Db      2821 AAGATTAAACAAGTGTGCTGAGTAATTTCACTAAATATGCTGTGAATTTTGTGATCAA
Qy      1979 -----1978

```

```

Db      2881 CTGTAGACAGAAATGTAATAATTTCACTCTCAAACTTCTGTTTGAATTAAGATGATAG
Qy      1979 -----T 1979
Db      2941 AGATTGCTTATGATGCTGCTTGTGCTCAACTTTTCTTCTGATGATTTTCTTTTCAATTT
Qy      1980 AGGCTAGCTAGAGCTGACATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATGATGAT
Db      3001 AGGCTAGCTAGAGCTGACATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATGATGAT
Qy      2040 GAATATATGATCCAGTGGCAGAAATTAATGCTCTGCTGCTTGGGCTGATGATGATGATGAT
Db      3061 GAATATATGATCCAGTGGCAGAAATTAATGCTCTGCTGCTTGGGCTGATGATGATGATGAT
Qy      2100 AGAATGTTACAG-----2113
Db      3121 AGAATGTTACAGAGGTGAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Qy      2114 -----AGTTTGGATGAGC 2128
Db      3181 TGGAATGATTAATAGTCTGCTGCTGCTTGTGATGATGATGATGATGATGATGATGATGAT
Qy      2129 GAATGCTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      3241 GAATGCTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      2189 ATTATACATGTTGAAATCTATCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT
Db      3301 ATTATACATGTTGAAATCTATCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      2249 CTCTGTTGGAAGCACTCTGAGAGAGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT
Db      3361 CTCTGTTGGAAGCACTCTGAGAGAGCTGCTGCTGATGATGATGATGATGATGATGATGATGAT
Qy      2309 ACAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      3421 ACAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      2369 GGTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2406
Db      3481 GGTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 3518

RESULT 7
US-10-600-070-10
; Sequence 10, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Koshmarova, Olga A.
; TITLE OF INVENTION: Placit Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-600-070-10

Query Match      71.5%; Score 1720.8; DB 7; Length 3667;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
Qy      1 ATGGAAGCTCTGAGTCAAGTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Db      481 ATGGAAGCTCTGAGTCAAGTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

```


QY 61 CGGCGAGCAAAAGCTCCGAGTAGCCACAACCTTACATATCTGCTCCGCGACG 120
DB 541 CGGCGAGCAAAAGCTCCGAGTAGCCACAACCTTACATATCTGCTCCGCGACG 600
QY 121 AATGAGGCGGAGCTCTTCTCTCCGACTTCAATTTTACCTCGGATTCCTCTCTCC 180
DB 601 AATGAGGCGGAGCTCTTCTCTCCGACTTCAATTTTACCTCGGATTCCTCTCTCC 660
QY 181 TTGCGCACCGCCACACACCGGCACTCTGCTCTCTGCGACCATCTATTTGATGCC 240
DB 661 TTGCGCACCGCCACACACCGGCACTCTGCTCTCTGCGACCATCTATTTGATGCC 720
QY 241 GAAACGCAAGTCCCATCCCATTTGATTTCTACAGGATTTAGAGCTCAACATTTTC 300
DB 721 GAAACGCAAGTCCCATCCCATTTGATTTCTACAGGATTTAGAGCTCAACATTTTC 780
QY 301 TTAAACGATGAAATCAAGAGCAATTCGAGCTTAGGCTTTCGAAACCGCGCAATTCG 360
DB 781 TTAAACGATGAAATCAAGAGCAATTCGAGCTTAGGCTTTCGAAACCGCGCAATTCG 840
QY 361 TTCAAGCGAGCAGCTTTTATCAGCCGAGACAGATTTCTTCAAGCTGTTGCCAACTTC 420
DB 841 TTCAAGCGAGCAGCTTTTATCAGCCGAGACAGATTTCTTCAAGCTGTTGCCAACTTC 900
QY 421 TCTAATCTCGGCTGAAGAGAGTACAAATGAAGCTCTTCTGATGATGAAGAGCTACA 480
DB 901 TCTAATCTCGGCTGAAGAGAGTACAAATGAAGCTCTTCTGATGATGAAGAGCTACA 960
QY 481 GTCATCACTGATGCTCTGGGAT----- 504
DB 961 GTCATCACTGATGCTCTGGGATTAAGTATTCGATTTCCGAAATTAAGTTCTTC 1020
QY 505 -----AAGTCTCTGGGCG 518
DB 1021 GTTTAATTTCAATGAATTTGATTAAGAGAGAACTTTTATCTAGTGAAGCTCTCGGAGC 1080
QY 519 TCTCTGTATTCAGAGAGAGTGTGAGCTGAGTATGTTCTTCGAGGTGTGAGGCTCT 578
DB 1081 TCTCTGTATTCAGAGAGAGTGTGAGCTGAGTATGTTCTTCGAGGTGTGAGGCTCT 1140
QY 579 GCTTAAGAGAGGTGCTTAAGCTGTTTAAAGCAAGATGTGTTTAACTTAAAGCGCTTGC 638
DB 1141 GCTTAAGAGAGGTGCTTAAGCTGTTTAAAGCAAGATGTGTTTAACTTAAAGCGCTTGC 1200
QY 639 GTTTCGATGCTCGAGGAGTGTATGCAATTTGATTCACCTGATTTTATCTGCTTA 698
DB 1201 GTTTCGATGCTCGAGGAGTGTATGCAATTTGATTCACCTGATTTTATCTGCTTA 1260
QY 699 TGAGTTTGTAGAGAGCTTGAAGCTTTTAC----- 730
DB 1261 TGAGTTTGTAGAGAGCTTGAAGCTTTTACAGGTAAGTTTGACTTCTGCTTAATTTG 1320
QY 731 -----A 731
DB 1321 ACGAGCTTGGCTTTATAAGAACTTCTGATTTGATTTGATTTGATTTGATTTGATTT 1380
QY 732 GAGAGAGAGAGCAATGACCTTTCGACCGGATTTACGTGCACAAATTTGATGAGCTTTG 791
DB 1381 GAGAGAGAGAGCAATGACCTTTCGACCGGATTTACGTGCACAAATTTGATGAGCTTTG 1440
QY 792 AGAGATACATCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGTATGATTTACGC 851
DB 1441 AGAGATACATCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGTATGATTTACGC 1500
QY 852 TGCGAAAAAGCTAATGATTTAAGCGGTGCGGAAATTTTGTGCTCTGTTGAGAGAG 911
DB 1501 TGCGAAAAAGCTAATGATTTAAGCGGTGCGGAAATTTTGTGCTCTGTTGAGAGAG 1560
QY 912 TGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGATTTTAAATGAGCGGTTTT 971
DB 1561 TGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGATTTTAAATGAGCGGTTTT 1620
QY 972 ACGAATGACAGCTGCTAGC----- 991

DB 1621 ATGAATGACAGCTGCTGAGCAGGTATACAGTTTAACTTTTAAATTTCTTTAGCA 1680
QY 992 -----AAGTTGATCTTTTGTGA 1008
DB 1681 TGAATATTAATTAGGTTTCTCATTTTAAATGATTTGTGTGTGTGTGTGTGTGTGT 1740
QY 1009 GCTAACCCCAAGCAATTTTCACAGAGCTCATTTGAAGTTTACGAAAGTTGCACTTCTCT 1068
DB 1741 GCTAACCCCAAGCAATTTTCACAGAGCTCATTTGAAGTTTACGAAAGTTGCACTTCTCT 1800
QY 1069 GTGGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGAGTCTATTAAGCAATTC 1128
DB 1801 GTGGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGAGTCTATTAAGCAATTC 1860
QY 1129 CAGCACTTCACAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACA 1188
DB 1861 CAGCACTTCACAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACA 1920
QY 1189 CGGAATTAATTTGGAGATGATCTCGCTTGAAGAGGGACTCTGCACTGCTTATAGGC 1248
DB 1921 CGGAATTAATTTGGAGATGATCTCGCTTGAAGAGGGACTCTGCACTGCTTATAGGC 1980
QY 1249 AAAAGTTGATGAATCCGTAATGTGTTGGCTTGAAGAGATTTCAATATATAGGAAT 1308
DB 1981 AAAAGTTGATGAATCCGTAATGTGTTGGCTTGAAGAGATTTCAATATATAGGAAT 2040
QY 1309 CCAAGTAATGTGAGGTTGTTTGTGAGAAATTCGAATCGATGATCAATGATGATCTCCCT 1368
DB 2041 CCAAGTAATGTGAGGTTGTTTGTGAGAAATTCGAATCGATGATCAATGATGATCTCCCT 2100
QY 1369 GGAATATGCAAAATTTGTGGAACCTGTGTGCGAGGGGTTGTCTTCTAGGTTCAAGAG 1428
DB 2101 GGAATATGCAAAATTTGTGGAACCTGTGTGCGAGGGGTTGTCTTCTAGGTTCAAGAG 2160
QY 1429 ACCAATGATTAATAATTTTAACTCGGGGACTATGATGATCTATGTTTGAATTC 1488
DB 2161 ACCAATGATTAATAATTTTAACTCGGGGACTATGATGATCTATGTTTGAATTC 2220
QY 1489 TTGGAAGAGTGAAGGTGATGTCAGGGTTTCTCTTATGCTGCTGCTCAACTATGCAAG 1548
DB 2221 TTGGAAGAGTGAAGGTGATGTCAGGGTTTCTCTTATGCTGCTGCTCAACTATGCAAG 2280
QY 1549 ATTGAGCCGAGCATGTGAAGCTAGTGTATGCAAGCACTGCAAGAAAGTTTCTCTCC 1608
DB 2281 ATTGAGCCGAGCATGTGAAGCTAGTGTATGCAAGCACTGCAAGAAAGTTTCTCTCC 2340
QY 1609 CGCTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGTA 1668
DB 2341 CGCTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGTA 2400
QY 1669 GATCCTGTGTGTAACAATGATGAGCGGTGATGAGCTGTGCTTATTTGCAAGAGCT 1728
DB 2401 GATCCTGTGTGTGTAACAATGATGAGCGGTGATGAGCTGTGCTTATTTGCAAGAGCT 2460
QY 1729 GTAAGACCTCTGTGAACCTTTGAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAG 1788
DB 2461 GTAAGACCTCTGTGAACCTTTGAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAG 2520
QY 1789 AGTAGCGTTGATGAACCTACTGTTGAATGTCCGTGCTGATATGTTTAAAGAGGCAAGT 1848
DB 2521 AGTAGCGTTGATGAACCTACTGTTGAATGTCCGTGCTGATATGTTTAAAGAGGCAAGT 2580
QY 1849 GTGAAGATCTACTGCTGTGTGTGGAATTTGCACTGATTTCACTGTTCAAGCAAGATAT 1908
DB 2581 GTGAAGATCTACTGCTGTGTGTGGAATTTGCACTGATTTCACTGTTCAAGCAAGATAT 2640
QY 1909 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGGATATGTTTCTTCTATGGAATCTGAT 1968
DB 2641 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGGATATGTTTCTTCTATGGAATCTGAT 2700
QY 1969 GTGCTTACCA----- 1978


```

Db      2701 GTCGCTACCATAGTATGATTAATGATGCAATTTTCATATATCTGCATTGCTCAATA 2760
Qy      1979 ----- 1978
Db      2761 TGCCTGTTTGTGAGCTAAGAACATAGTCCCACTTAATACATGTCCCAAAAGTTGTACC 2820
Qy      1979 ----- 1978
Db      2821 AAGATTAAACAAGTTGCTGAATTAATTTCACTAATTAATGCTGTAATTTTTCATCAAA 2880
Qy      1979 ----- 1978
Db      2881 CTGTAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTAGAATAGTAGATTAAG 2940
Qy      1979 ----- 1979
Db      2941 AGATTGCTTAGTGTGCTTGTGCAACTTTTCTTCTGTGATTTTCTTTTCATTT 3000
Qy      1980 AGGCTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGA 2039
Db      3001 AGGCTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGA 3060
Qy      2040 GAAATATAGTATCCAAAGTGGCAGAAAGATTAAGTCTTGTGGCTTTTGGGCTGATCACCGCAT 2099
Db      3061 GAAATATAGTATCCAAAGTGGCAGAAAGATTAAGTCTTGTGGCTTTTGGGCTGATCACCGCAT 3120
Qy      2100 AGAAATTTTACCAG----- 2113
Db      3121 AGAAATGTTACAGAGGTGAGGGAATTAATCTACAAATTCATCAATTTGTGAAACCTGT 3180
Qy      2114 -----AGTTTGTGATGGG 2128
Db      3181 TGCAATGATTAATAGTGTGCTGCTTGTGATTTCTGTATTAATAGTTTGTGATGGG 3240
Qy      2129 GAATGCTGAAGATTTGAGCTGACAGACAGCTGAACTGGCAGCTTGGTGTGTTATG 2188
Db      3241 GAATGCTGAAGATTTGAGCTGACAGACAGCTGAACTGGCAGCTTGGTGTGTTATG 3300
Qy      2189 AATTATACATGTTGAACTATCTGTGTCAGAGTGCAGAGTCTCAGCAGATGGAACCCGCTG 2248
Db      3301 AATTATACATGTTGAACTATCTGTGTCAGAGTGCAGAGTCTCAGCAGATGGAACCCGCTG 3360
Qy      2249 CTCTGCTGGAAGCACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTCATCAGAAA 2308
Db      3361 CTCTGCTGGAAGCACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTCATCAGAAA 3420
Qy      2309 ACAATGCTACTGATGTCAGAACCTACACAACAAGATACGAAGTTTCTGTGTCAGAGTCA 2368
Db      3421 ACAATGCTACTGATGTCAGAACCTACACAACAAGATACGAAGTTTCTGTGTCAGAGTCA 3480
Qy      2369 GGTGGAATCACTGAAGGCTCTGTTCTTGCATCATTA 2406
Db      3481 GGTGGAATCACTGAAGGCTCTGTTCTTGCATCATTA 3518

```

RESULT 8

```

US-10-600-070-132/c
; Sequence 132, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 561

```

```

; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (127)..(127)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (520)..(520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-132

```

```

Query Match 20.0%; Score 482; DB 7; Length 561;
Best Local Similarity 99.4%; Pred. No. 1e-134;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1922 GCTCATCTTTTCAACCGAAGATTAAGTCTGCTCTATGGAATCTGATGCTGCTACGCTAG 1981
Db      561 GCTCATCTTTTCAACCGAAGATTAAGTCTGCTCTATGGAATCTGATGCTGCTACGCTAG 502
Qy      1982 GGTCAAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGAGA 2041
Db      501 GGTCAAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGAGA 442
Qy      2042 AATATAGTATCCAAAGTGGCAGAAAGATTAAGTCTGCTTGTGGCTTGTGATCACCGCAT 2101
Db      441 AATATAGTATCCAAAGTGGCAGAAAGATTAAGTCTGCTTGTGGCTTGTGATCACCGCAT 382
Qy      2102 AATATGTTACCAAGCTTTTGGATGAGGCGAATGCTGGAAGATTTTGGATGCAAGAGCCTG 2161
Db      381 AATATGTTACCAAGCTTTTGGATGAGGCGAATGCTGGAAGATTTTGGATGCAAGAGCCTG 322
Qy      2162 AAATGCGCAGCTTGGGTTGTTATGATTAATCACTGTGAAATCTATCTGTGACAGTG 2221
Db      321 AAATGCGCAGCTTGGGTTGTTATGATTAATCACTGTGAAATCTATCTGTGACAGTG 262
Qy      2222 TGAAGTCTGAGCAGATGGAACCCGCTGCTGTGTGGAAGCACTTGGAGAGTCTGCTT 2281
Db      261 TGAAGTCTGAGCAGATGGAACCCGCTGCTGTGTGGAAGCACTTGGAGAGTCTGCTT 202
Qy      2282 GTCTATCTGATTTGTTGTTTCAATCCAGAAACAAATGCTATGATGTCAGAACTTACACAA 2341
Db      201 GTCTATCTGATTTGTTGTTTCAATCCAGAAACAAATGCTATGATGTCAGAACTTACACAA 142
Qy      2342 GATACGAAGTTTCTGGTCCAGATCAGGAGGGAATCACTGAAGGCTCTGTTCTTGCAT 2401
Db      141 GATACGAAGTTTCTGGTCCAGATCAGGAGGGAATCACTGAAGGCTCTGTTCTTGCAT 82
Qy      2402 CATTA 2406
Db      81 CATTA 77

```

RESULT 9

```

US-10-600-070-126
; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2

```

SEQ ID NO 126
LENGTH: 2283
TYPE: DNA
ORGANISM: *Oryza sativa*
US-10-600-070-126

Query Match 20.0%; Score 481; DB 7; Length 2283;
Best Local Similarity 54.3%; Pred. No. 5.2e-134;
Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

108 CTGCTCCGCGCAGCAATATGGCCGACCGCTTCTCTCCGACTTCAATTTCACTTCCATTC 167
132 CTGGCGCGAAGCGCTCTTCCGCACTTCCACCTTCCCAACCGCGCCGCTCCGACCC 191
168 CTCTCTCTCTCTCTGCGCACCGGCACCAACCGGACCTCTGCTCTTCCGACCATC 227
192 GCGCTCCCGGCGCGCGCGCGCGCGCGCGCGCTCCGCTCCCGCTTCCGCTCTT 251
228 TATTGATCTGCTCCGAAAGCGCAGCTCCCATCCCATTTGATTTCTACAGGATTTAGAGC 287
252 CCGCGAGCGCGCGAAGCTCTCCCTCCGCTCCAAAGTGAATTTCTACAGGATTTAGAGG 311
288 TCAAAACATTTCTTAAACGATGGAATCAGAGAGCATTTGAAAGCTAGGCTTCAAAAC 347
312 AGAGCAATTTCTTCTGCGCATGAGCATCAGAGGCGCTTCAGAGCAAGATAGCAAGCC 371
348 GCGGCAATTCGCTTTCAGGAGCAGCGCTTAAATCAGCGCGGAGCAAGATTTCAAGCTGC 407
372 ACCGCAATATGCTACAGCAGAGATGCTTGTGTGCTGCGACAAATATGCTGCAAGATTGC 431
408 TTGCGAACTCTGCTCTATCTCGCTCTAGAGAGATCAATAGAGCTCTTCTTGAATGA 467
432 CCAATGACCTCTATGAGACAGAACTCCGCACTCAAGTATGATGCTGCGCTTCTGAGAA 491
468 TGAAGAAGCTACAGTCACTGATGTTCTTGGGATAGAGTTCCTGCGCTCTCTGTGT 527
492 CCGTGAAGAAGCTCTCACTCAATGATATGCTTGGGACAGAGAGCTTGGG----- 540
528 ATTGCAAGAAGTGTGAGACTGAGATGTTCTTCCGCTTGTGAGAGCTCTGCTTAAAGA 587
541 -----GAGGCACTTGTGCTTGTGTAATCTGAGAGAAAGTTGCTTCTGGA 584
588 GAGGTTCCCTAAGTCGTTTAAAGCAAGATGCTTTTAAATGAGCGCTCTGCTTCTCGA 647
585 TCGGCAACCCAGCGCTTCAAGAGAGAGCTGTGCTGATGCGATGCTCTGCTTATATGGA 644
648 TGTCTCGAGGATGCTATGAGCATTTGATCACTGATTTTATTAATCTGATATGAGATTGT 707
645 TCTATCAAGGAGTGTCTATGAGCAAGCCCTCCAGATGTAATGCTGCTGCAAGGTGCT 704
708 TGAAGAAGCTTTGAAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
705 CGAGAGGCTCTCAAGCTCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 764
768 TGAAGAAATGATGAGACTTTGAGAGAGATCACTCCGCTTATATGCTTTGAGACTATGAG 827
765 TTCAACAGATGAGAACTCTCAAGAGATTTACACTCGCTGTATTTGAGAGCTCTCTC 824
828 CTTACCGCTTGTGATGATTAACGCTGCGAAAGACTAAATGCTTTAAAGCGGTGTGCGAA 887
825 CTTTCATTTGACACAGAGATCATTAAGAGCGCCAGAGAGGCTTCAAGGTGTGAGAA 884
888 TATTTTGTGCTGTGAGAGAGATGAGCATCAAGCTCTTGTGTGGGCTTTGACCGCTGA 947
885 CATTTGTGAGGCTGTGAGAGAGAGATGATGCTTACCGTTGAGAGAGATTTTCTCTGA 944
948 GAAGTTATGATGAGGCGTTTAAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007
945 AGCTTCATTAAGAGGCTTTTGTGAGAGATGATCAATTTGAAGAGATGAGATTTCTTTC 1004
1008 AGCTACCCCAAGCAATATTCAGAGAGATCAATTTGAAGTTTACAGAGTTGACATTTGCT 1067
1005 AAAAAACCGAATAGCATTTCTCTGATGATGTTGAATTTTACATTTGACATTTGACA 1064

1068 TGTGCTCAAGCTTTTATGTTAGAGAGCCACCTTTTACAGAGATGCTGATAGCAATT 1127
1065 TGTGCTCAAGCAATTTATAGTAAAGGCCAATTCATCATGATGCGGATGATCTTTT 1124
1128 CCAAGCACTTCCAGAGCTTAAGGTAAATGCTATGAGATTTCTGCGCATTTTATGATAC 1187
1125 TGAACAACTCCAGAACTTCA-----CATAGTTCTCATTA 1160
1188 ACGAATTAATTTGGAGATAGACTTGGCTTAAAGAGGAGACTCTGACACTCTTATAG 1247
1161 TGTATATGATATGATATGATGACTTGCATTTGAAAGGAGATTTGCTCATTTGCTAGT 1220
1248 CAAAGTTATGATAGCGGATGCTGTTGGGCTTAAACAGTGAAGATTTCAATATAGAA 1307
1221 AGATTTACAGATGAGATGAGATGCTTGAATTTGATATGATGCTTCCATACAGAGA 1280
1308 TCCAGCTATTTGAGAGTTGTTTGGAGATTTCAATC--GTGATGACAAATGATATCT 1364
1281 CCCCAAAATTTAGAGTTTATGTTGACCAACTCTAGCATCAGTGAAGAAATGATCTCT 1340
1365 CCGTGAATATGCAATTTGTGAAACCTGTTTGGCAGGGGTTGCTTCTTCTAGTTGAG 1424
1341 TCCAGGCGCTGTGCAAGCTTTTGAAGCTTGGCTTATCTTTGAGTTTCTTATGAGAG 1400
1425 AGACACCAAGATTAATAATTTAACTCGGAGACTATATGATGATCTATGCTTATGAG 1484
1401 AGATATCGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1460
1485 TTACTTGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1544
1461 CTACCTTGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1520
1545 AAGATTTGAGCGCGAGCATGTAAGAGCTAGTGTATGAGAGGACCTGACAGAAATTTTCC 1604
1521 AAACTGTGTGCTCAAGCTACAGCTGCACTTGTAT----- 1555
1605 TTCCGCTTATACAGATAGAAACTCGCTGAACCCAGAGATGTCAGAGAGACAGTGTAG 1664
1556 -----CTGTGAATCAAAATGCTATTCAGAGCG----- 1581
1665 TGTAGATCTGTTGTAACAATGTAAGCCGTATGATGAGCCTGTGCTTATTTGCAAG 1724
1582 -----TTCAACAA 1589
1725 AGCTGTAAGCCCTCTGAAACCTTTGAATCTATATGATTTATGCAATTCAGAGTGGCTTC 1784
1590 GGTTTTTCATGATGATGAACAGTTAGACAGTCAAGCCATGAGAAATATCTAAAGATGCCC 1649
1785 AGAGATGAGCGTTGATGAACACTGTTGAATATGCTGTTGCTGATATGTTAAAGAGGC 1844
1650 TGGGGATATCTTGAATAATTTTGAACAGAGAAATGACCTGCTCATGATTCGAGAAATG 1709
1845 AAGTGAAGATCTAGCTGTGTGTGTCGAATGAGATGATTTTCACTGTTACGCCA 1904
1710 CGCTTGAAGATATCTGTGTGTGCGCACTGTTGATGAGATGAGTGTGAGAGAGAGAG 1769
1905 GTATTTTCTTAAAGAGCTCATCTTTCAACCAAGATATGATGTTCTTCTATGGAATC 1964
1770 ATATTTGCTCTGTAAGAGGCCCTTTCTGCTATTTAGAGTGAAGAGATGATGCTGCACT 1829
1965 TGAT-----GTGCTACCAATAGGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2015
1830 TGTATTAATGATGATCTTATGATATCTGCTGATGATGATGATGATGATGATGATGATG 1889
2016 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2075
1890 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1949
2076 GGTCTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2135
1950 GGCCTTGGGACAGAACTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2009


```

; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
; US-10-600-070-135

```

```

Query Match      11.6%; Score 278.8; DB 7; Length 660;
Best Local Similarity 67.3%; Pred. No. 5.2e-73;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;

```

```

QY 81 ACCTACCCACACACCTCTACAACTATCTGCTCCGACGAAATGGGCGGACCTCTTCT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 ACCTAACCGTCTCATTTCTCCGCGCTCTCCGACCAAGTAATGGGCGGACGACTCAT 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 CTCGCACTTCATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 TTCGATTTCCATTTCTCCGCGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 201 CGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 AGTCACTCTC-----ACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 CATTGATTTCTACAGGATTTAGAGAGCTCAACACATTTCTTAACGATGAATTCAGAA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 TCTGCACTCTGTAACAAATCTCCGCGCCGAAAGCATTCTCTGCGATGATTTGCGAG 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 AGCATTCGAAGCTAGGATTTGAAACCGCGCAATTCGATTCGATTCGACGACGCTTTAT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 AGCTTATGAAGCGAAATTCGAAAGCTCTCTCAATGATTTGATGATGAACTTTGAT 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 CAGCCGGAACAGATTTCTTCAAGCTGCTGCGAACTCTCTCTCTCTCTCTCTCTCTCT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 TACTCTCTCTCAATTTCTTCAAGCTGCTGTAAGAACCTTCACTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 AGATTCGAATGAAGCTCTCTCTGATGATGAA-----GAAGCTACAGTCACTAC 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 AGGATTAATCAAAAGCTCTCTCAAGATGAAGAAAGATGAGAAATCTTCACTCTCAC 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 TGAATTTCTTGGGATTAAGGTTCTCTGAGGCTCTCTGTAATTCAGAAAGATGAGAG 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 TGAATTCCTTTGACAAAGTTCTCTGAGCTCTCTGCTGCTGCTGCTGCTGCTGAGAG 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 TGAATTAAGTTCTTGGGATTTGAGAGCTCTCTGTAATGAAGAGATTCCTTAAGTCTTAA 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 GGAATTTGCTCTCGAATTTGAGAGGAGGTTACTGAGAGAGAGTTACCGAAGATGTTAA 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 GGAAGATTTGCTTTGATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 GCAAGATTTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 ATTGATTCACCTGATTTTATTAAGTCTGATTAAGATTTGATGA 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 611 TTTGTCCCGCCGAGATTTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

; US-10-424-599-35059
; Sequence 35059, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 35059
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_131660C.1
; US-10-424-599-35059

```

```

Query Match      9.6%; Score 231.6; DB 7; Length 1039;
Best Local Similarity 64.4%; Pred. No. 1.5e-58;
Matches 380; Conservative 0; Mismatches 204; Indels 6; Gaps 2;

```

```

QY 1823 TTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTTGAC 1882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 50 TTACTGAAGAAATTAACATGCAAGTGTGCAATGATCATGATGCTGTGTGATTTGAGAC 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1883 TGATTTCACTGTTCAACGAGATATTTTCTTAAAGCA---GCTCATCTTTCAACGCA 1939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 TGGTAATCTTGTGTTGTTTAAAGTTTAACTTAACTTAACTTAACTTAACTTAACT 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1940 AGGATATGTTCTCTTATGGAATCTGATGTGCTCAATAGGCTCAGTCAAGCTGACG 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170 AATGACTGTTCTCGCAATGATGTGCAATGATCTAATCAATTAATGTTCTTGGAGATGAG 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2000 ATTCAAGAG---CACTTCCAGATGATGATGATGATGATGATGATGATGATGATGAT 2056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 AAAATGAGAGAGCTTACCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2057 GGCAGAAATTAAGTCTCTGCTTTTGGGCTGTGATCAACGATGAAATGTTTACAGAG 2116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 290 GGCAGAAATGTTAAATCCGAAGCTTTTGGACCTGACCACTTGTGGAGAGGTTGCATGAG 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2117 TTTTGGATGGGGAATGCTGAAGATTTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAG 2176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 TTTTGAACGATGAGATGTTGAGATTAAGATGATGATGATGATGATGATGATGATGATGAT 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2177 GGTGATTTATGATTAATTAACATGTTGAACTATCTGTTGACAGTGTGACAGTCTCAG 2236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 410 GTTGTCTATGATCTACCTTGAAGAGCTCTCAATGATGATGATGATGATGATGATGATGAT 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2237 ATGAAACCGTCTCTGTTGAAAGCACTCTGAGAGAGTCTGTTCTATCTGATTTG 2296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 470 ATGGGCGCGTGTGAGTGTGAAACAACTGTAAGAGTCTACTACCTGAATGCGGTAG 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2297 TTATTCAGAAAAAATGCTATGATGTGAGAACTTACACAAATGATGAGATTTTCT 2356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 530 GTCATCCGAAACAGATGCTTTTAATGAGAACTTACACAAATGATGAGATTTTCT 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2357 GGTCAAGTCAAGGAGTGAATATCACTGAAGGCTCTGTTCTTGATCATTA 2406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 590 TTAAGAGCGGAGATGAGAAATTTGTAAGAGAGCTCTCTTGAAGTCTTA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

```

; US-10-600-070-143
; Sequence 143, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oesteryoung, Katherine W.
; APPLICANT: Vitha, Stanislaw
; APPLICANT: Kosharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206

```

; SOFTWARE: Patentin version 3.2
; SEQ ID NO 143
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-600-070-143

Query Match 9.3%; Score 224; DB 7; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.9e-56;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```
QY 555 AGTCTTCGGGTTGTGAGGCTCTGCTTAAGAGAGGTGCTTAAGTGGTTTAAGCAAGA 614
    |||||
DB 2 AGTCTTGCAATTGAGGGCACTTACTGAGAGACCGCCGCCAAGCGTTCAAGCAGGA 61
    |||||
QY 615 TGTGTTTATGATGCGGCTTCCGTTCTCGATGTCGAGAGGATGCTATGCAATTGGA 674
    |||||
DB 62 TGTGTGCTGGCAATGCGCTCGCTTAATGATCTATCAAGGACGCAATGCGGCTAG 121
    |||||
QY 675 TCCACCTGATTTTATTACTGTATGAGTTGTTGAGAGACTTTGAACTTTTACAGA 734
    |||||
DB 122 CCTCCAGATGTAATCGCTGCTGTGAGTGTGAAAGGCTCTCAAGCTTTTGACAGA 181
    |||||
QY 735 GGAAGAGCAAGTAGACCTTGACCGAATTTACGTGCAAAATTGATGAGACTTTGAAAGA 794
    |||||
DB 182 GGATGGGGCAATCAATCTGCACCTGGTTGCTCTCAAAATTGATGAACCTGAGAGA 241
    |||||
QY 795 GATCACTCCGCGTTATGTCCTTGAGACTTACCGCTTGGTATGATTAAGCTGC 854
    |||||
DB 242 TATCACACCTCGTGTGTTTGAAGCTTCTGCCCCCTTCCTGATGAAACATCAGAA 301
    |||||
QY 855 GAAAGACTAATAGTTTAAAGCGGTGCGAATATTTTGTGCTGTTGAGAGAGTGG 914
    |||||
DB 302 TGAACACCAAGAGTCTTGTGTGTGAGAAACATTTTGTGAGTGTGCAAGAGAGG 361
    |||||
QY 915 AGCATCAGCTCTTGTGGGGTTTGAACCGGTGAGAAATTATGATGAGGCGTTTTCAG 974
    |||||
DB 362 TATTGTAAGTGTGAGAGATTTTGGGTGAAGCTACATGATGAAAGCTTCTGCA 421
    |||||
QY 975 AATGACAGCTGCTGAGCAGGTTATCTTTTGTAGCTAACCCCAAGCAATATTCAGCAGA 1034
    |||||
DB 422 GATGACATGCGCGAGAGATGATTTCTCTCAAAAACACGAAATAGCATACCGCTGA 481
    |||||
QY 1035 GTCAATTGAAAGTTAAGAGTGAAGTGAAGTGTGCTTGTGAGCTCAAGCTTTTATGGA 1090
    |||||
DB 482 ATGCTTGAATCTAATAGCGTGGCACTTGCAAAATGTTGCTCAAGCAATTTGAAGTA 537
    |||||
```

Search completed: February 21, 2006, 15:09:20
Job time : 2718 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 14:04:13 ; Search time 545 Seconds
(without alignments)
7847.364 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagctctgagtcacgt.....gtctgtcttcgcatcataa 2406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodaca/1/ina/1 COMB.seq: *
2: /cgn2_6/ptodaca/1/ina/5 COMB.seq: *
3: /cgn2_6/ptodaca/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodaca/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodaca/1/ina/H COMB.seq: *
6: /cgn2_6/ptodaca/1/ina/PTUS COMB.seq: *
7: /cgn2_6/ptodaca/1/ina/PP COMB.seq: *
8: /cgn2_6/ptodaca/1/ina/RE COMB.seq: *
9: /cgn2_6/ptodaca/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.6	2.4	7218	2 US-08-232-463-14	Sequence 14, Appl
2	50.2	2.1	57280	3 US-09-949-016-11796	Sequence 11796, A
3	50.2	2.1	57280	3 US-09-949-016-12843	Sequence 12843, A
4	50.2	2.1	57280	3 US-09-949-016-12844	Sequence 12844, A
5	50.2	2.1	57280	3 US-09-949-016-12846	Sequence 12846, A
6	50.2	2.1	57280	3 US-09-949-016-13542	Sequence 13542, A
7	50.2	2.1	57280	3 US-09-949-016-13543	Sequence 13543, A
8	50.2	2.1	57280	3 US-09-949-016-13544	Sequence 13544, A
9	50.2	2.1	57280	3 US-09-949-016-13545	Sequence 13545, A
10	50.2	2.1	57280	3 US-09-949-016-14633	Sequence 14633, A
11	50.2	2.1	57280	3 US-09-949-016-14634	Sequence 14634, A
12	50.2	2.1	57280	3 US-09-949-016-14635	Sequence 14635, A
13	50.2	2.1	57280	3 US-09-949-016-14636	Sequence 14636, A
14	50.2	2.1	57280	3 US-09-949-016-14637	Sequence 14637, A
15	50.2	2.1	57280	3 US-09-949-016-14638	Sequence 14638, A
16	50.2	2.1	57280	3 US-09-949-016-14639	Sequence 14639, A
17	50.2	2.1	57280	3 US-09-949-016-14640	Sequence 14640, A
18	49.4	2.1	913	2 US-08-217-327-3	Sequence 3, Appl1
19	49.4	2.1	913	2 US-07-885-870A-3	Sequence 3, Appl1
20	49.4	2.1	913	2 US-08-298-687A-3	Sequence 3, Appl1
21	49.4	2.1	913	2 US-08-530-797-2	Sequence 2, Appl1
22	49.4	2.1	913	2 US-08-298-829-3	Sequence 3, Appl1
23	49.4	2.1	913	2 US-08-787-335-2	Sequence 2, Appl1
24	49.4	2.1	1984	2 US-07-885-970A-25	Sequence 25, Appl1

25	49.4	2.1	1985	2 US-08-298-687A-25	Sequence 25, Appl1
26	49.4	2.1	1985	2 US-08-298-829-25	Sequence 25, Appl1
27	49.2	2.0	1334	3 US-09-547-693-234	Sequence 234, Appl
28	48.4	2.0	152132	3 US-09-949-016-13845	Sequence 13845, A
29	48.4	2.0	152145	3 US-09-949-016-12371	Sequence 12371, A
30	46.8	1.9	114793	3 US-10-148-806-3	Sequence 3, Appl1
31	44.8	1.9	832	3 US-09-621-976-2813	Sequence 2813, Appl
32	44.2	1.8	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
33	44.2	1.8	4411529	3 US-09-103-840A-1	Sequence 2, Appl1
34	44	1.8	1141	3 US-09-806-708B-22	Sequence 22, Appl
35	43	1.8	402	3 US-10-131-827-8166	Sequence 8166, Appl
36	43	1.8	1281	3 US-09-991-181-415	Sequence 415, Appl
37	43	1.8	1281	3 US-09-990-444-415	Sequence 415, Appl
38	43	1.8	1281	3 US-09-997-333-415	Sequence 415, Appl
39	43	1.8	1281	3 US-09-992-598-415	Sequence 415, Appl
40	43	1.8	3265	3 US-09-832-129-113	Sequence 13, Appl1
41	42.8	1.8	60990	3 US-09-949-016-14080	Sequence 14080, A
42	42.4	1.8	39154	3 US-09-949-016-12384	Sequence 12384, A
43	42.4	1.8	39154	3 US-09-949-016-12801	Sequence 12801, A
44	42.4	1.8	39443	3 US-09-949-016-14326	Sequence 14326, A
45	42.4	1.8	39443	3 US-09-949-016-14327	Sequence 14327, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s

Db 12921 CATCATCAGCA 12911

Query Match	2.1%;	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.0059;		

Db 13041 ACCATCAATATCCCTTACGTCACCAACCGCTTTCACATCAACCAACCAACCAACC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGTCCCAAGCCACGCTCCCATCCCAT 264
Db 12981 ACAACCAACCACTCATATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 11
US-09-949-016-14634/c

; Sequence 14634, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14634

; LENGTH: 57280

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-14634

Query Match

Best Local Similarity 53.9%; Score 50.2; DB 3; Length 57280;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGCAGCAAAATGGCGACCGCTCTCTCC 144
Db 13101 AGCACCATACCAACCAACCTGCTCCACCAACCATCAACCTTACCATTAAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTTATGACTTCCACCATCAACCGCTTTCACATCAACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGTCCGAAGCCACGCTCCCATCCCAT 264
Db 12981 ACAACCAACCACTCATATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 12

US-09-949-016-14635/c

; Sequence 14635, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14635
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14635

Query Match

Best Local Similarity 53.9%; Score 50.2; DB 3; Length 57280;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGCAGCAAAATGGCGACCGCTCTCTCC 144
Db 13101 AGCACCATACCAACCAACCTGCTCCACCAACCATCAACCTTACCATTAAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTTATGACTTCCACCATCAACCGCTTTCACATCAACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGTCCGAAGCCACGCTCCCATCCCAT 264
Db 12981 ACAACCAACCACTCATATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 13

US-09-949-016-14636/c

; Sequence 14636, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14636

; LENGTH: 57280

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-14636

Query Match

Best Local Similarity 53.9%; Score 50.2; DB 3; Length 57280;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGCAGCAAAATGGCGACCGCTCTCTCC 144
Db 13101 AGCACCATACCAACCAACCTGCTCCACCAACCATCAACCTTACCATTAAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTTATGACTTCCACCATCAACCGCTTTCACATCAACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGTCCGAAGCCACGCTCCCATCCCAT 264
Db 12981 ACAACCAACCACTCATATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 13:35:30 : Search time 1421 Seconds
(without alignments)
11284.483 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagcctcgtacacgt.....gtctgtctctcgtacataa 2406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406	100.0	2406	12	ADJ38129 Arabidops
2	2402.8	99.9	2406	12	ADJ38135 Arabidops
3	2399.6	99.7	2406	12	ADJ38208 Plaetid d
4	2399.6	99.7	2637	12	ADJ38210 Plaetid d
5	2399.6	99.7	2679	13	ADJ14901 Plant CDN
6	1724	71.7	3667	12	ADJ38130 Arabidops
7	1720.8	71.5	3667	12	ADJ38136 Arabidops
8	482	20.0	561	12	ADJ38212 Plaetid d
9	481	20.0	2283	12	ADJ38206 Plaetid d
10	328.8	13.7	631	12	ADJ38264 Plaetid d
11	278.8	11.6	660	12	ADJ38215 Plaetid d
12	224	9.3	537	12	ADJ38223 Plaetid d
13	222.4	9.2	552	13	ACN48855 Cotton pr
14	207	8.6	545	12	ADJ38255 Plaetid d
15	197.2	8.2	491	12	ADJ38254 Plaetid d
16	195.8	8.1	647	12	ADJ38265 Plaetid d
17	193.8	8.1	460	12	ADJ38267 Plaetid d
18	190	7.9	446	12	ADJ38231 Plaetid d
19	186.6	7.8	652	12	ADJ38266 Plaetid d

20	178.6	7.4	607	13	ACN48781
21	175	7.3	439	12	ADP93143
22	170.4	7.1	608	12	ADJ38217
23	170.2	7.1	563	12	ADJ38258
24	169.8	7.1	622	12	ADJ38226
25	165.4	6.9	527	12	ADJ38214
26	158.6	6.6	871	12	ADJ38232
27	155.4	6.5	307	12	ADJ38218
28	154.656	6.4	161	12	ACH86330
29	153.056	6.4	509	13	ACN59803
30	153.056	6.4	2355	13	ADJ14689
31	152.256	6.3	1880	14	ADW18179
32	151.872	6.3	366	12	ADN99165
33	151.656	6.3	1504	13	ADJ31985
34	151.264	6.3	1470	13	ADN83960
35	151.272	6.3	1671	13	ADJ36390
36	151.272	6.3	3245	13	ADJ27315
37	151.072	6.3	3290	13	ACN39669
38	150.856	6.3	429	12	ADJ30054
39	150.856	6.3	4068	13	ADV40896
40	150.656	6.3	923	13	ADJ49034
41	150.656	6.3	1117	13	ADJ65456
42	150.656	6.3	1389	13	ADJ44953
43	150.656	6.3	2760	13	ADV41745
44	150.464	6.3	909	12	ADN99181
45	150.472	6.3	1356	12	ADN01090

ALIGNMENTS

RESULT 1	ADJ38129	standard; cDNA; 2406 BP.
ID	ADJ38129	
XX	ADJ38129;	
XX	06-MAY-2004	(first entry)
DT	06-MAY-2004	(first entry)
DE	Arabidopsis thaliana Arc6-1 cDNA Sequ1.	
XX	prokaryotic type; plastid division; Fuz2; ARC6; ARCS; Fzo; plant cell;	
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;	
KW	herbicide target; gene; ss.	
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX	WO2004001003-A2.	
PN	WO2004001003-A2.	
XX	31-DEC-2003.	
PD	31-DEC-2003.	
XX	20-JUN-2003; 2003WO-US019536.	
PF	20-JUN-2003; 2003WO-US019536.	
XX	20-JUN-2002; 2002US-0390140P.	
PR	09-AUG-2002; 2002US-0402242P.	
PR	20-JUN-2003; 2003US-00600070.	
XX	(UNMS) UNIV MICHIGAN STATE.	
PA	(UNMS) UNIV MICHIGAN STATE.	
XX	Osteryoung KM, Vittha S, Koksharova OA, Gao H;	
PI	Osteryoung KM, Vittha S, Koksharova OA, Gao H;	
XX	WPI, 2004-082486/08.	
DR	WPI, 2004-082486/08.	
DR	P-PSDB; ADJ38202.	
XX	P-PSDB; ADJ38202.	
PT	New isolated Fuz2, ARCS and/or Fzo-like nucleic acid sequences, useful	
PT	for further characterizing plastid division in plant cells, and in	
PT	varying agronomic and horticultural characteristics of economically	
PT	important plants.	
XX	Claim 1, SEQ ID NO 1; 287bp; English.	
PS	Claim 1, SEQ ID NO 1; 287bp; English.	
XX	This invention relates to novel prokaryotic type or plastid division and	
CC	related genes and proteins. In particular, the invention relates to novel	

ACN48781	Cotton pr
ADP93143	Cotton ex
ADJ38217	Plaetid d
ADJ38258	Plaetid d
ADJ38226	Plaetid d
ADJ38214	Plaetid d
ADJ38232	Plaetid d
ADJ38218	Plaetid d
ACH86330	Human gen
ACN59803	Cotton gy
ADJ14689	Plant CDN
ADW18179	Pinus rad
ADN99165	Novel hum
ADJ31985	Plant ful
ADN83960	Plant ful
ADJ36390	Plant ful
ADJ27315	Plant ful
ACN39669	Tumour-as
ADJ30054	3' end of
ADV40896	Rat card
ADJ49034	Plant ful
ADJ65456	Plant ful
ADJ44953	Bacterial
ADV41745	Rat card
ADN99181	Novel hum
ADN01090	Streptoco

CC Fun2 (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;

Query Match 100.0%; Score 2406; DB 12; Length 2406;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGAAGCTCTGATGATCGATCGGCAATTTGCTCTCCCAATTCATATAGCCGATTAACA 60
DB 1 ATGGAAGCTCTGATGATCGATCGGCAATTTGCTCTCCCAATTCATATAGCCGATTAACA 60
QY 61 CCGGCGACGCAAAAGCTCCGACGTAGACCAACAACCTCTACACTATCTGCTCCGACG 120
DB 61 CCGGCGACGCAAAAGCTCCGACGTAGACCAACAACCTCTACACTATCTGCTCCGACG 120
QY 121 AATAGGCGCGACCGCTCTTCTCTCGACTTCAATTTCACTCGAATTCCTCTCTCTCC 180
DB 121 AATAGGCGCGACCGCTCTTCTCTCGACTTCAATTTCACTCGAATTCCTCTCTCTCC 180
QY 181 TTGCGCACCGCCACACACCGGCACTCTGTCTCTCTGCGACCACTATATGATCGTCC 240
DB 181 TTGCGCACCGCCACACACCGGCACTCTGTCTCTCTGCGACCACTATATGATCGTCC 240
QY 241 GAAGCGCAGCTCCCAATCCCATTTGATTTCTACACAGATTTAGAGAGCTCAACAATTC 300
DB 241 GAAGCGCAGCTCCCAATCCCATTTGATTTCTACACAGATTTAGAGAGCTCAACAATTC 300
QY 301 TTAAACGATGGAATCAGAAAGCAATTCGAAGCTTTCGAAACCGCGCAATTCGAT 360
DB 301 TTAAACGATGGAATCAGAAAGCAATTCGAAGCTTTCGAAACCGCGCAATTCGAT 360
QY 361 TTTCAGGAGAGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGTTCGAAACTCTG 420
DB 361 TTTCAGGAGAGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGTTCGAAACTCTG 420
QY 421 TCTAATCTCGGCTGAAAGAGATCAATGAAGGTCTTCTGATGATGAAGAAGCTACA 480
DB 421 TCTAATCTCGGCTGAAAGAGATCAATGAAGGTCTTCTGATGATGAAGAAGCTACA 480
QY 481 GTCATCACTGATGTTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGAAGAAGGT 540
DB 481 GTCATCACTGATGTTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGAAGAAGGT 540
QY 541 GGTGAGACTGAGATTAAGTTCTTGGGTTGTTGAGGCTTCTGCTTAAGAGAGGTTGCTTAAG 600
DB 541 GGTGAGACTGAGATTAAGTTCTTGGGTTGTTGAGGCTTCTGCTTAAGAGAGGTTGCTTAAG 600
QY 601 TCGGTTAAGCAAGATGTTGTTTATGATTAAGGCTTGCATTCGATGCTGAGAGGAT 660
DB 601 TCGGTTAAGCAAGATGTTGTTTATGATTAAGGCTTGCATTCGATGCTGAGAGGAT 660
QY 661 GCTATGCGCATTTGATCACTGATTTTATTAAGGCTTGCATTCGATGCTGAGAGGAT 720
DB 661 GCTATGCGCATTTGATCACTGATTTTATTAAGGCTTGCATTCGATGCTGAGAGGAT 720
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGCACCGGATTTTACGTCACAATTTGAT 780
DB 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGCACCGGATTTTACGTCACAATTTGAT 780
QY 781 GAAACTTTGAGAAAGATCACTCCGCTTATGTTGAGGCTACTGAGCTTAACGCTTGT 840
DB 781 GAAACTTTGAGAAAGATCACTCCGCTTATGTTGAGGCTACTGAGCTTAACGCTTGT 840
QY 841 GATGATTAAGCTGAGAAAGACTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCT 900
DB 841 GATGATTAAGCTGAGAAAGACTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCT 900
```

```
QY 901 GTTGGAGAGGTGAGACATCACTCTTGTGGGGGTTTGACCCGAGAGAGTTTATGAT 960
DB 901 GTTGGAGAGGTGAGACATCACTCTTGTGGGGGTTTGACCCGAGAGAGTTTATGAT 960
QY 961 GAGGCGTTTTCAGAAATGACAGCTGTGAGCAGTTGATCTTTTGTGACTACCCCAAGC 1020
DB 961 GAGGCGTTTTCAGAAATGACAGCTGTGAGCAGTTGATCTTTTGTGACTACCCCAAGC 1020
QY 1021 AATATTCACAGAGATCATTTTAAGTTTACGAAGTTTGACCTTGTGCTCAAGCT 1080
DB 1021 AATATTCACAGAGATCATTTTAAGTTTACGAAGTTTGACCTTGTGCTCAAGCT 1080
QY 1081 TTTATTTGTAAGAGCACAACCTTTTACAGATGCTGATTAACAAATTCAGCAACTTCAG 1140
DB 1081 TTTATTTGTAAGAGCACAACCTTTTACAGATGCTGATTAACAAATTCAGCAACTTCAG 1140
QY 1141 CAGGCTAAGGTAATGCTATGAGATCTCGGAGTGTGTATGATACAGGAATATTTGG 1200
DB 1141 CAGGCTAAGGTAATGCTATGAGATCTCGGAGTGTGTATGATACAGGAATATTTGG 1200
QY 1201 GAGATTAAGCTTGTGCTTAAAGGGGACTGTGCACTGCTTATAGGCAAGTTGATGA 1260
DB 1201 GAGATTAAGCTTGTGCTTAAAGGGGACTGTGCACTGCTTATAGGCAAGTTGATGA 1260
QY 1261 TGCCGTAATGTTGGCTTAAAGAGTGAAGATTCACAAATATAGGAATCAGCTATTTG 1320
DB 1261 TGCCGTAATGTTGGCTTAAAGAGTGAAGATTCACAAATATAGGAATCAGCTATTTG 1320
QY 1321 GAGTTGTTTGAAGATTCAAATGCTGATGACATGATGATCTCCCTGACATATGCA 1380
DB 1321 GAGTTGTTTGAAGATTCAAATGCTGATGACATGATGATCTCCCTGACATATGCA 1380
QY 1381 TTGTTGGAACCTGTTGGCAGGGGTGCTTCCAGGTTAGAGACCAAAAGATTA 1440
DB 1381 TTGTTGGAACCTGTTGGCAGGGGTGCTTCCAGGTTAGAGACCAAAAGATTA 1440
QY 1441 AATTTAACTCGGGGACTACTATGATGATTCATATGTTTGAAGTTACTTGAAGAGTG 1500
DB 1441 AATTTAACTCGGGGACTACTATGATGATTCATATGTTTGAAGTTACTTGAAGAGTG 1500
QY 1501 GAGGTAGTTCAAGGTTCTCTTAACTGCTGCAACTATGAGCAAGATTTGAGACCGAG 1560
DB 1501 GAGGTAGTTCAAGGTTCTCTTAACTGCTGCAACTATGAGCAAGATTTGAGACCGAG 1560
QY 1561 CATGTAAAGCTATGCTATGACAGGACCTGCAAGAAAGTTTCTCTCCGCTATACAGAT 1620
DB 1561 CATGTAAAGCTATGCTATGACAGGACCTGCAAGAAAGTTTCTCTCCGCTATACAGAT 1620
QY 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATAGTATCCTGTTGGT 1680
DB 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATAGTATCCTGTTGGT 1680
QY 1681 AACATGTAAGCCGATAGGTGAGCTGTGTCTTATTTAGCAAGACTGTAAAGCCTCT 1740
DB 1681 AACATGTAAGCCGATAGGTGAGCTGTGTCTTATTTAGCAAGACTGTAAAGCCTCT 1740
QY 1741 GAAAACCTTGAACCTAATGATTAATGCAATTCAGAGCTGGGGTCTCAAGAGTACGTTGAT 1800
DB 1741 GAAAACCTTGAACCTAATGATTAATGCAATTCAGAGCTGGGGTCTCAAGAGTACGTTGAT 1800
QY 1801 GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGTGAATCTTA 1860
DB 1801 GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGTGAATCTTA 1860
QY 1861 GCTGTGTGTGCAATTTGACTGATTTTCACTGTTCAGGCAAGATATTTTCTTAAAGC 1920
DB 1861 GCTGTGTGTGCAATTTGACTGATTTTCACTGTTCAGGCAAGATATTTTCTTAAAGC 1920
QY 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATAGGAATCTGATGCTGCTACCATTA 1980
DB 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATAGGAATCTGATGCTGCTACCATTA 1980
```


Db 841 GATGATAGCTGCGAAAAGACTAAATGCTTTAAGCGGTGCGGAAATTTTGTGCTC 900
 Qy 901 GTTGGAGAGGTGAGCAATCAGCTCTTTGTTGGGGTTTGAACCCGTGAGAGTTTATGAT 960
 Db 901 GTTGGAGAGGTGAGCAATCAGCTCTTTGTTGGGGTTTGAACCCGTGAGAGTTTATGAT 960
 Qy 961 GAGGCGTTTATGCAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAC 1020
 Db 961 GAGGCGTTTATGCAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAC 1020
 Qy 1021 AATATTCAGCAGAGCTCATTGAGATTGAAGTTTGAAGTTGCACTTGCTTGTGCTCAAGCT 1080
 Db 1021 AATATTCAGCAGAGCTCATTGAGATTGAAGTTTGAAGTTGCACTTGCTTGTGCTCAAGCT 1080
 Qy 1081 TTTATTTGGTAAGACCACTTTTACAGATGCTGATTAAGAAATTTCCAGCAACTTCAG 1140
 Db 1081 TTTATTTGGTAAGACCACTTTTACAGATGCTGATTAAGAAATTTCCAGCAACTTCAG 1140
 Qy 1141 GAGGCTAAGGTATGCTATGAGATTCCTGCGATGTTGATTAACAAGCAATTAATTTGG 1200
 Db 1141 GAGGCTAAGGTATGCTATGAGATTCCTGCGATGTTGATTAACAAGCAATTAATTTGG 1200
 Qy 1201 GAGATAGACTTGGCTCTAGAAAGGGGACTCTGTGCACTGCTTATAGCAAAAGTTGATGA 1260
 Db 1201 GAGATAGACTTGGCTCTAGAAAGGGGACTCTGTGCACTGCTTATAGCAAAAGTTGATGA 1260
 Qy 1261 TGCCTGATGTGTTGGGCTTTAGACAGTGAAGATTCAATATAGAAATCAAGTATTTGG 1320
 Db 1261 TGCCTGATGTGTTGGGCTTTAGACAGTGAAGATTCAATATAGAAATCAAGTATTTGG 1320
 Qy 1321 GAGTTTGTGTTGAGAAATCAATGCTGATGACAAATGATGATCTCCCTGCACTATGCAAA 1380
 Db 1321 GAGTTTGTGTTGAGAAATCAATGCTGATGACAAATGATGATCTCCCTGCACTATGCAAA 1380
 Qy 1381 TTTTGGAAAACCTGTTGGCAGGGTGTCTTTCTTAGTTTCAGAGACACCAAGATAAA 1440
 Db 1381 TTTTGGAAAACCTGTTGGCAGGGTGTCTTTCTTAGTTTCAGAGACACCAAGATAAA 1440
 Qy 1441 AAATTTAACTCGGGGACTATATGATGATCTTATGCTTTGAGTTTCTTGAAGAGTGTG 1500
 Db 1441 AAATTTAACTCGGGGACTATATGATGATCTTATGCTTTGAGTTTCTTGAAGAGTGTG 1500
 Qy 1501 GAGGTAGTTCAAGGCTCTCTTAGCTGCTGCAAGTATGAGCAAGATTTGGAGCCGAG 1560
 Db 1501 GAGGTAGTTCAAGGCTCTCTTAGCTGCTGCAAGTATGAGCAAGATTTGGAGCCGAG 1560
 Qy 1561 CATGTGAAGCTAGTCTATGACAGGCACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
 Db 1561 CATGTGAAGCTAGTCTATGACAGGCACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
 Qy 1621 AGAAACTCGGCTGAAACCAAGAGTGTGCAAGAGAGCTGTTATGTTAGTCTGTTGGT 1680
 Db 1621 AGAAACTCGGCTGAAACCAAGAGTGTGCAAGAGAGCTGTTATGTTAGTCTGTTGGT 1680
 Qy 1681 AACCATGTAGCCCTGATGCTGAGCCCTGCTGCTTTATGCAAACTGTAAACCCCTCT 1740
 Db 1681 AACCATGTAGCCCTGATGCTGAGCCCTGCTGCTTTATGCAAACTGTAAACCCCTCT 1740
 Qy 1741 GAAAACTTTGAAACTATATGATTAATGCAATTCGAGCTGCGGCTCTCAGAGAGTACGTTGAT 1800
 Db 1741 GAAAACTTTGAAACTATATGATTAATGCAATTCGAGCTGCGGCTCTCAGAGAGTACGTTGAT 1800
 Qy 1801 GAAACTACTGTGAAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
 Db 1801 GAAACTACTGTGAAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
 Qy 1861 GCTGCTGCTGCTGCAATTTGACCTGTTACAGCCAGAGATTTTCTTAAAGC 1920
 Db 1861 GCTGCTGCTGCTGCAATTTGACCTGTTACAGCCAGAGATTTTCTTAAAGC 1920
 Qy 1921 AGCTATCTTTTCAAGCAAGATGATGTTCTTCTATGGAATCTGATGCTGATCAAT 1980
 Db 1921 AGCTATCTTTTCAAGCAAGATGATGTTCTTCTATGGAATCTGATGCTGATCAAT 1980

Db 1921 AGCTATCTTTTCAAGCAAGATGATGTTCTTCTATGGAATCTGATGCTGATCAAT 1980
 Qy 1981 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGATGATGATGACCTGACAG 2040
 Db 1981 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGATGATGATGACCTGACAG 2040
 Qy 2041 AATATAGTATCCAGATGTCAGAAATTAATGCTCTGCTTTTGGGCTGATTCACCGCA 2100
 Db 2041 AATATAGTATCCAGATGTCAGAAATTAATGCTCTGCTTTTGGGCTGATTCACCGCA 2100
 Qy 2101 GAAATGTTACAGAGGTTTGGATGGGAAATGCTGAAGATTGACAGACAGCAAGCT 2160
 Db 2101 GAAATGTTACAGAGGTTTGGATGGGAAATGCTGAAGATTGACAGACAGCAAGCT 2160
 Qy 2161 GAAACTGCGCAGCTTGGGTTTATGATTAATGACCTGTTGAACTATCTGTTGACAGT 2220
 Db 2161 GAAACTGCGCAGCTTGGGTTTATGATTAATGACCTGTTGAACTATCTGTTGACAGT 2220
 Qy 2221 GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGTGGAAGCAACTCTGAGAGCTGCT 2280
 Db 2221 GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGTGGAAGCAACTCTGAGAGCTGCT 2280
 Qy 2281 TGTCTATCTGATTTGCTTCAATCAGAAACCAATGCTATGATGAGAACCTTACCA 2340
 Db 2281 TGTCTATCTGATTTGCTTCAATCAGAAACCAATGCTATGATGAGAACCTTACCA 2340
 Qy 2341 AGATACGAAGTTTCTGCTCAGAGTCAGGTCGAAATCACTGAAGGCTCTGTTCTTGA 2400
 Db 2341 AGATACGAAGTTTCTGCTCAGAGTCAGGTCGAAATCACTGAAGGCTCTGTTCTTGA 2400
 Qy 2401 TCATPA 2406
 Db 2401 TCATPA 2406
 RESULT 3
 ADJ38208
 ID ADJ38208 strand; DNA; 2406 BP.
 XX
 AC ADJ38208;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plasmid division-related Arc6 orthologue gene 2.
 XX
 KW prokaryotic type; plasmid division; Fnu2; ARCS; Fzo; plant cell;
 XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; gene; ds.
 OS Arabidopsis thaliana.
 XX
 PN W02004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WC-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 XX
 PR 09-AUG-2002; 2002US-0402242P.
 XX
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oesteryoung KW, Vithe S, Kosharova OA, Gao H;
 XX
 DR MPI; 2004-082486/08.
 XX
 DR P-PSDB; ADJ38207.
 XX
 PT New isolated Fnu2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plasmid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.

PS Disclosure; Fig 8; 287bp; English.

XX This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fnu2 (Ara6), APCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 12; Length 2406;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGAAGCTCTGATGATCGTGGCAATGATGCTCTCCCATTCGAATATGCGATTCACA 60
DB 1 ATGAAGCTCTGATGATCGTGGCAATGATGCTCTCCCATTCGAATATGCGATTCACA 60
QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCCACAACTCTAACAATATCTGCTCCGACG 120
DB 61 CCGGCGACGACAAAGCTCCGACGTAAGCCACAACTCTAACAATATCTGCTCCGACG 120
QY 121 AATATGGGCGGACCGCTCTTCTCTCCGATTCATTTTCACTCTCCGATTCCTCTCTCC 180
DB 121 AATATGGGCGGACCGCTCTTCTCTCCGATTCATTTTCACTCTCCGATTCCTCTCTCC 180
QY 181 TTGCGCACGCGCACACACGCGCACTCTGCTCTGCGACCATATATGATGATGCGC 240
DB 181 TTGCGCACGCGCACACACGCGCACTCTGCTCTGCGACCATATATGATGATGCGC 240
QY 241 GAAAGCGACGCTCCCATCCCATGATTTTCTACAGGTAATGAGAGCTCAACAATTC 300
DB 241 GAAAGCGACGCTCCCATCCCATGATTTTCTACAGGTAATGAGAGCTCAACAATTC 300
QY 301 TTAAACCGATGAGATTCAGAAAGCATTCGAACTGAGGTTTCGAAACCGCGCAATTCG 360
DB 301 TTAAACCGATGAGATTCAGAAAGCATTCGAACTGAGGTTTCGAAACCGCGCAATTCG 360
QY 361 TTCAAGGACGACGCTTTAATCAGCGGAGACAGATTTCTCAAGCTGCTCGAAATCTG 420
DB 361 TTCAAGGACGACGCTTTAATCAGCGGAGACAGATTTCTCAAGCTGCTCGAAATCTG 420
QY 421 TCTAATCTCTGCTCTAGAAAGAGTACATGAAAGTCTTCTGATGATGAAAGAGCTACA 480
DB 421 TCTAATCTCTGCTCTAGAAAGAGTACATGAAAGTCTTCTGATGATGAAAGAGCTACA 480
QY 481 GTTATCATGATGATTTCTTGGATTAAGTCTTGGGCTCTCTGTGATATGCAAGAGT 540
DB 481 GTTATCATGATGATTTCTTGGATTAAGTCTTGGGCTCTCTGTGATATGCAAGAGT 540
QY 541 GGTGAGATGAGATGATTTCTTGGGTTGATGAGGCTCTGCTGATGATGCAAGAGT 600
DB 541 GGTGAGATGAGATGATTTCTTGGGTTGATGAGGCTCTGCTGATGATGCAAGAGT 600
QY 601 TCGTTTAAAGCAAGATGATTTTAACTGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TCGTTTAAAGCAAGATGATTTTAACTGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GCTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GCTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 AAGCTTTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 721 AAGCTTTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 GAGACTTTGAGAGAGATCACTCGCGTATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GAGACTTTGAGAGAGATCACTCGCGTATGATGATGATGATGATGATGATGATGATGAT 840
```

```
QY 841 GATGATTAAGCTCGAAGAAAGCTAAATGTTTAAAGCGGTGCGGAATATTTGTGCT 900
DB 841 GATGATTAAGCTCGAAGAAAGCTAAATGTTTAAAGCGGTGCGGAATATTTGTGCT 900
QY 901 GTTGAAGAGGTGAGGATCAGCTCTTGTGAGGAGTTTGAACCGGTGAAGATTATGAT 960
DB 901 GTTGAAGAGGTGAGGATCAGCTCTTGTGAGGAGTTTGAACCGGTGAAGATTATGAT 960
QY 961 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCAGC 1020
DB 961 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCAGC 1020
QY 1021 AATATTCAGCAGATCATTTAAGTTTACGAATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATATTCAGCAGATCATTTAAGTTTACGAATGATGATGATGATGATGATGATGAT 1080
QY 1081 TTTATTTGTAAGAGCACCTTTTACAGATGCTGATAGCAATTCAGCAATTCAG 1140
DB 1081 TTTATTTGTAAGAGCACCTTTTACAGATGCTGATAGCAATTCAGCAATTCAG 1140
QY 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
DB 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
QY 1201 GAGATGACCTTGGCTAGAAAGGAGACCTGCGACTGCTTAPAGGCAAGTTGATGAA 1260
DB 1201 GAGATGACCTTGGCTAGAAAGGAGACCTGCGACTGCTTAPAGGCAAGTTGATGAA 1260
QY 1261 TGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 TGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGACAAATGATGATGATGATGATGATGAT 1380
DB 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGACAAATGATGATGATGATGATGATGAT 1380
QY 1381 TTGTTGAAACCTGTTTGGAGGAGTGTCTTCTTCTAGGTTGAGAGACACCAAGATAA 1440
DB 1381 TTGTTGAAACCTGTTTGGAGGAGTGTCTTCTTCTAGGTTGAGAGACACCAAGATAA 1440
QY 1441 AATATTTAACTCGGAGACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 AATATTTAACTCGGAGACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 CATGTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 CATGTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATGATGATGATGATGATGAT 1680
DB 1621 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATGATGATGATGATGATGAT 1680
QY 1681 AACATGTAAGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 AACATGTAAGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 GAAAACCTTGAACCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 GAAAACCTTGAACCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GAAACTACTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GAAACTACTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1861 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
```

Oy	1921	GGCTCATCTTTTCAACGCAAGATATGGTTCCTCTATAGGAATCGANATGTCGTACATA	1980
Db	1921	AGCTCATCTTTTCAACGCAAGATATGGTTCCTCTATAGGAATCGANATGTCGTACATA	1980
Oy	1981	GGGTCAGTCAACAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTCGTAGACTGCAGAG	2040
Db	1981	GGGTCAGTCAACAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTCGTAGACTGCAGAG	2040
Oy	2041	AATATAGTATCCAAAGTGGCAGAGATTAAAGTCTCTGGCTTTTGGGCTGTATCACCGCATTA	2100
Db	2041	AATATAGTATCCAAAGTGGCAGAGATTAAAGTCTCTGGCTTTTGGGCTGTATCACCGCATTA	2100
Oy	2101	GAATATGTTACCAAGAGCTTTTGGATGGGGAAATGCTGAAAGATTGGAATGACAGAGCAAGT	2160
Db	2101	GAATATGTTACCAAGAGCTTTTGGATGGGGAAATGCTGAAAGATTGGAATGACAGAGCAAGT	2160
Oy	2161	GAATCTGGCAGCTTGGGTTTATGATTAATACACTGTGTAATCTATCTGTGACAGT	2220
Db	2161	GAATCTGGCAGCTTGGGTTTATGATTAATACACTGTGTAATCTATCTGTGACAGT	2220
Oy	2221	GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGTGGAAAGCACTCTGAGAGAGTCTGCT	2280
Db	2221	GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGTGGAAAGCACTCTGAGAGAGTCTGCT	2280
Oy	2281	TGTCATCTGATTTGGTTCATCCAGAAAACAATGTAATGATGTCAGAACTTACACAA	2340
Db	2281	TGTCATCTGATTTGGTTCATCCAGAAAACAATGTAATGATGTCAGAACTTACACAA	2340
Oy	2341	AGATACGAAGTTTCTGGTCCAAAGTCAGGGTGGAAAATCACTGAAGGCTCTGTTCTTGCA	2400
Db	2341	AGATACGAAGTTTCTGGTCCAAAGTCAGGGTGGAAAATCACTGAAGGCTCTGTTCTTGCA	2400
Oy	2401	TCATTA 2406	
Db	2401	TCATTA 2406	
RESULT 4			
ADJ38210			
ID	ADJ38210	standard; DNA; 2637 BP.	
AC	ADJ38210;		
DT	06-MAY-2004	(first entry)	
DE	Plastid division-related <i>Arc6</i> orthologue gene 3.		
KM	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;		
KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;		
KW	herbicide target; gene; ds.		
OS	Arabidopsis thaliana.		
PN	WO2004001003-A2.		
XX	31-DEC-2003.		
XX	20-JUN-2003; 2003WO-US019536.		
XX	20-JUN-2002; 2002US-0390140P.		
PR	09-AUG-2002; 2002US-0402242P.		
PR	20-JUN-2003; 2003US-00600070.		
PA	(UNMS) UNIV MICHIGAN STATE.		
XX	OsterYoung KM, Vitha S, Koksharova OA, Gao H;		
XX	WPI; 2004-082486/08.		
XX	P-PSDB; ADJ38209.		
PT	New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful		
PT	for further characterizing plastid division in plant cells, and in		
PT	varying agronomic and horticultural characteristics of economically		

Query Match	Best Local Similarity	Score	DB	Length
Matches 2402; Conservative	99.7%; Pred. No. 0;	2399.6;	12;	2637;
	0;	Mismatches	4;	Indels
	0;	Gaps	0;	
Sequence 2637 BP, 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;				
1	ATGGAAGCTCTGAGTCAAGTCAAGTGGTCTCTCCCATTCGATTAAGCGATTACCA	60		
114	ATGGAAGCTCTGAGTCAAGTGGTCTCTCCCATTCGATTAAGCGATTACCA	173		
61	CCGGGACGACAAAGCTCCGACGTAGCACAACCTCTACACATATCTGCTCCGACG	120		
174	CCGGGACGACAAAGCTCCGACGTAGCACAACCTCTACACATATCTGCTCCGACG	233		
121	AAATGGGCGACCGCTCTCTCCGACCTTAATTTCACTTCGATTTCTCTCTCTCC	180		
234	AAATGGGCGACCGCTCTCTCTCCGACCTTAATTTCACTTCGATTTCTCTCTCTCC	293		
181	TTTCGCAACCGCACCAACACCGGCACTCTCGTCTCTCGGACCATTAATGATGTCGC	240		
294	TTTCGCAACCGCACCAACACCGGCACTCTCGTCTCTCGGACCATTAATGATGTCGC	353		
241	GAAGCGACGTCCCATCCCATTTGATTTCTACACAGTATTAGAGCTCAAAACATTTTC	300		
354	GAAGCGACGTCCCATCCCATTTGATTTCTACACAGTATTAGAGCTCAAAACATTTTC	413		
301	TTAACCGATGATATCAGAAAGACATTTGCAAGTGAAGTTTCGAAACCGCCGCAATTCG	360		
414	TTAACCGATGATATCAGAAAGACATTTGCAAGTGAAGTTTCGAAACCGCCGCAATTCG	473		
361	TTACGCGACGACGCTTTAATCAGCCGAGACAGATTTCTTCAAGCTGCTTGGAACTCG	420		
474	TTACGCGACGACGCTTTAATCAGCCGAGACAGATTTCTTCAAGCTGCTTGGAACTCG	533		
421	TCTATATCCCGGTCTAGAAAGAGATGACATGAAGGCTCTTCTGATGATGAAGACTACA	480		
534	TCTATATCCCGGTCTAGAAAGAGATGACATGAAGGCTCTTCTGATGATGAAGACTACA	593		
481	GTCATCATGTAGTCTCTTGGGATTAAGGTTCTCTGGGAGCTCTCTGTGTGATTCGAAAGGT	540		
594	GTCATCATGTAGTCTCTTGGGATTAAGGTTCTCTGGGAGCTCTCTGTGTGATTCGAAAGGT	653		
541	GGTGAGACTGAGATATGTTCTTGGGTTGGTGAAGGCTCTGCTTAAAGAGAGTTGCTTAG	600		
654	GGTGAGACTGAGATATGTTCTTGGGTTGGTGAAGGCTCTGCTTAAAGAGAGTTGCTTAG	713		
601	TGCTTTAAGCAAGATGGTATTAGTAAAGGCGTCTGCGTTCTGATGCTCGAGAGAT	660		
714	TGCTTTAAGCAAGATGGTATTAGTAAAGGCGTCTGCGTTCTGATGCTCGAGAGAT	773		
661	GCTATGAGATTGATTCACCTGATTTTATACGTGTTATGATTTGTTGAGAGCTTTG	720		
774	GCTATGAGATTGATTCACCTGATTTTATACGTGTTATGATTTGTTGAGAGCTTTG	833		
721	AAGCTTTTACAGAGAGAAAGACAAATGACCTTGCACCGGATTTACGTGACCAATTTAT	780		
834	AAGCTTTTACAGAGAGAAAGACAAATGACCTTGCACCGGATTTACGTGACCAATTTAT	893		
781	GAGACTTTGGAAGATACATCCCGGTTATGCTTGGAGCTACTTGGCTTACCGGCTGAT	840		

Db 894 GAGACTTTGGAAAGATCACTCCGCTTATGCTTTGAGACTTATGCTTACCGCTTGT 953
 Qy 841 GATGATTAACCTGCGAAGAACTTAAATGCTTTAAGCGGTGCGCAATATTTTGTGCT 900
 Db 954 GATGATTAACCTGCGAAGAACTTAAATGCTTTAAGCGGTGCGCAATATTTTGTGCT 1013
 Qy 901 GTTGAAGAGGTGAGCATGCTCTTGTGGGGTGTGACCCGTGAGAAATTTATGAAT 960
 Db 1014 GTTGAAGAGGTGAGCATGCTCTTGTGGGGTGTGACCCGTGAGAAATTTATGAAT 1073
 Qy 961 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGTACCCCAAGC 1020
 Db 1074 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGTACCCCAAGC 1133
 Qy 1021 AATATTCACAGAGTCAATTTGAAGTTTGAAGTTGACCTTGTCTTGTGGCTCAAGCT 1080
 Db 1134 AATATTCACAGAGTCAATTTGAAGTTTGAAGTTGACCTTGTCTTGTGGCTCAAGCT 1193
 Qy 1081 TTTATTTGTAAGAACCACTTTTACAGGATGCTATTAAGCAATTCAGCACTTCAAG 1140
 Db 1194 TTTATTTGTAAGAACCACTTTTACAGGATGCTATTAAGCAATTCAGCACTTCAAG 1253
 Qy 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATATATACAGCAATTAATTTG 1200
 Db 1254 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATATATACAGCAATTAATTTG 1313
 Qy 1201 GAGATGACCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTATAGCAAAATTTGATGA 1260
 Db 1314 GAGATGACCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTATAGCAAAATTTGATGA 1373
 Qy 1261 TGCCTATGATGTTGGGCTTGAAGATGAGGATTCACAAATTAAGCAATTCAGCAATTTG 1320
 Db 1374 TGCCTATGATGTTGGGCTTGAAGATGAGGATTCACAAATTAAGCAATTCAGCAATTTG 1433
 Qy 1321 GAGTTGTTTGGAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Db 1434 GAGTTGTTTGGAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
 Qy 1381 TGTGTTGAAAACCTGCTTGGCAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1440
 Db 1494 TGTGTTGAAAACCTGCTTGGCAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1553
 Qy 1441 AATTTTAACTCGGGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 1554 AATTTTAACTCGGGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
 Qy 1501 GAGGTAGTTCAAGGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 Db 1614 GAGGTAGTTCAAGGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1673
 Qy 1561 CATGTGAAGCTTATGCTATGACAGCACTGCAAGAAATTTTCTTCTTCTTCTTCTTCTTCT 1620
 Db 1674 CATGTGAAGCTTATGCTATGACAGCACTGCAAGAAATTTTCTTCTTCTTCTTCTTCTTCT 1733
 Qy 1621 AGAACTCGGCTGAACCAAGATGTCAGAGACAGTGTGTTAGTGAATCCTGTTGGT 1680
 Db 1734 AGAACTCGGCTGAACCAAGATGTCAGAGACAGTGTGTTAGTGAATCCTGTTGGT 1793
 Qy 1681 AACATGTAGCGCTGATGCTGAGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
 Db 1794 AACATGTAGCGCTGATGCTGAGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1853
 Qy 1741 GAAACTTTGAAACTATGATTAATGCAATTCAGCTGCGGTCTCAGAGATGAGCTTGTAT 1800
 Db 1854 GAAACTTTGAAACTATGATTAATGCAATTCAGCTGCGGTCTCAGAGATGAGCTTGTAT 1913
 Qy 1801 GAAACTTACTGTTGAATGCTGCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
 Db 1914 GAAACTTACTGTTGAATGCTGCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1973
 Qy 1861 GCT 1920

Db 1974 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2033
 Qy 1921 AGCTCATCTTTTCAAGCAAGATATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
 Db 2034 AGCTCATCTTTTCAAGCAAGATATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2093
 Qy 1981 GGGTCACTGCAAGCTGACATTCAGAGCACTTCCAGAAATGATGCTAGACTGCAAG 2040
 Db 2094 GGGTCACTGCAAGCTGACATTCAGAGCACTTCCAGAAATGATGCTAGACTGCAAG 2153
 Qy 2041 AATATGATTCAGAGTGCAGAAAGATTAAGTCTGCTTGGCTTGGCTGATCACCGATA 2100
 Db 2154 AATATGATTCAGAGTGCAGAAAGATTAAGTCTGCTTGGCTTGGCTGATCACCGATA 2213
 Qy 2101 GAAATGTTACAGAGTTTGAATGGCGCAATGCTGAAGTTTGAAGTGAAGTGAAGTGAAG 2160
 Db 2214 GAAATGTTACAGAGTTTGAATGGCGCAATGCTGAAGTTTGAAGTGAAGTGAAGTGAAG 2273
 Qy 2161 GAAACTGCGAGCTTGGGTTGTTTATGATTAATACATGTTGAAGTATCTGTTGACAGT 2220
 Db 2274 GAAACTGCGAGCTTGGGTTGTTTATGATTAATACATGTTGAAGTATCTGTTGACAGT 2333
 Qy 2221 GTGACAGTCTCAGCAAGATGGAACCTGCTCTGCTGAGAGCAACTGAGAGTCTGCT 2280
 Db 2334 GTGACAGTCTCAGCAAGATGGAACCTGCTCTGCTGAGAGCAACTGAGAGTCTGCT 2393
 Qy 2281 TGTCTATCTGATTTGGTTTCTTCAAGAAACAAATGCTATGATGCTGCAAGTCAACACA 2340
 Db 2394 TGTCTATCTGATTTGGTTTCTTCAAGAAACAAATGCTATGATGCTGCAAGTCAACACA 2453
 Qy 2341 AGATGCAAGTCTTCTGCTCAAGTCAAGGCTGGAATATCACTGAAGCTCTGTTCTTGA 2400
 Db 2454 AGATGCAAGTCTTCTGCTCAAGTCAAGGCTGGAATATCACTGAAGCTCTGTTCTTGA 2513
 Qy 2401 TCATTA 2406
 Db 2514 TCATTA 2519

RESULT 5
 ADT14901
 ID ADT14901 standard; cDNA; 2679 BP.
 XX
 AC ADT14901;
 DT 13-JAN-2005 (first entry)
 XX
 DE Plant cDNA, Seq ID 227.
 XX
 KW Plant; sb; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 OS Viridiplantae.
 PN US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 PF 18-DEC-2003; 2003US-00739930.
 XX
 PR 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 XX
 PA (KOVA/) KOVALIC D K.
 XX
 PI Kovalic DK;
 XX
 DR WPI; 2004-757369/74.

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetic, and in particular for producing transgenic plants with improved
PT biological characteristics.

XX Claim 1; SEQ ID NO 227; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPRO at
CC seqdata.uspro.gov/sequence.html?docid=20040216190.

XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;

XX Query Match 99.7%; Score 2399.6; DB 13; Length 2679;

XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTTGTGATCGTTCGATTCGCTCCCATTCGATTAAGCGATTACCA 60
DB 115 ATGGAAGCTTGTGATCGTTCGATTCGCTCCCATTCGATTAAGCGATTACCA 174
QY 61 CCGGCGACGACAAAGCTCCGACGTAGGCAACAACCTCTAACAATCTCTCGCAGC 120
DB 175 CCGGCGACGACAAAGCTCCGACGTAGGCAACAACCTCTAACAATCTCTCGCAGC 234
QY 121 AAATGGGCGACCGCTCTTCTCTCGACTTCGATTCGATTCGATTCGATTCGATTC 180
DB 235 AAATGGGCGACCGCTCTTCTCTCGACTTCGATTCGATTCGATTCGATTCGATTC 294
QY 181 TTGGCCACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCG 240
DB 295 TTGGCCACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCG 354
QY 241 GAAAGCGACGTCCTCCATCCCATTTGATTTTACCAAGGATTTAGAGCTCAACAATTC 300
DB 355 GAAAGCGACGTCCTCCATCCCATTTGATTTTACCAAGGATTTAGAGCTCAACAATTC 414
QY 301 TTAACCGATGGAATCAGAAAGGCAATTGGAAGCTAAGGTTTGAACCGCGCAATTGGT 360
DB 415 TTAACCGATGGAATCAGAAAGGCAATTGGAAGCTAAGGTTTGAACCGCGCAATTGGT 474

QY 361 TTCAGCGACGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTCGAACTCTG 420
DB 475 TTCAGCGACGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTCGAACTCTG 534
QY 421 TCTAATCTCCGCTCTGAAAGAGATTCATGAAAGCTCTCTTGAATGATGAAGAGCTAC 480
DB 535 TCTAATCTCCGCTCTGAAAGAGATTCATGAAAGCTCTCTTGAATGATGAAGAGCTAC 594
QY 481 GTCATCACTGATGTTCTTGGGATTAAGGTTCTCGGAGCTCTCTGATTTGCAAGAGGT 540
DB 595 GTCATCACTGATGTTCTTGGGATTAAGGTTCTCGGAGCTCTCTGATTTGCAAGAGGT 654
QY 541 GGTGACATGAGATAGTCTTCTCGGTTGATGAGCTCTGCTTAAGAGAGGTTGCTTAAG 600
DB 655 GGTGACATGAGATAGTCTTCTCGGTTGATGAGCTCTGCTTAAGAGAGGTTGCTTAAG 714
QY 601 TCGTTTAAGCAAGATGCTGTTTATGATGAGGCTCTGCTTCTGATGCTCGAGAGAT 660
DB 715 TCGTTTAAGCAAGATGCTGTTTATGATGAGGCTCTGCTTCTGATGCTCGAGAGAT 774
QY 661 GCTAAGGCAATGGATCAGCTGATTTTATTAATGATGATGATGATGATGATGATGAT 720
DB 775 GCTAAGGCAATGGATCAGCTGATTTTATTAATGATGATGATGATGATGATGATGAT 834
QY 721 AAGCTTTTACAG 780
DB 835 AAGCTTTTACAG 894
QY 781 GAGACTTTGAAAGAGATCTCGGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGT 840
DB 895 GAGACTTTGAAAGAGATCTCGGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGT 954
QY 841 GATGATTAAGCTGCGAAG 900
DB 955 GATGATTAAGCTGCGAAG 1014
QY 901 GTTGAAGAGAGTGAAGATCAGCTCTTGTGGGGTTTGAACCGTGAGAGATTATGAAT 960
DB 1015 GTTGAAGAGAGTGAAGATCAGCTCTTGTGGGGTTTGAACCGTGAGAGATTATGAAT 1074
QY 961 GAGGCGTTTATGAGATGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1075 GAGGCGTTTATGAGATGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
QY 1021 AATATTCAGAGAGAGATTTGAAGTTTGAAGTTGCACTTGTGAGCTCAAGCT 1080
DB 1135 AATATTCAGAGAGAGATTTGAAGTTTGAAGTTGCACTTGTGAGCTCAAGCT 1194
QY 1195 TTTATTTGTAAG 1254
QY 1141 CAGGCTAAGGTAATGAGCTTATGAGATTCCTGCGATTTGTATGATCAAGGAATATTTG 1200
DB 1255 CAGGCTAAGGTAATGAGCTTATGAGATTCCTGCGATTTGTATGATCAAGGAATATTTG 1314
QY 1201 GAGATGACCTTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1315 GAGATGACCTTGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1374
QY 1261 TGCCTATGATGTTGGGCTTGAACAGTGAAGATTCACATATAGAGATTCAGATATTTG 1320
DB 1375 TGCCTATGATGTTGGGCTTGAACAGTGAAGATTCAGATATAGAGATTCAGATATTTG 1434
QY 1321 GAGTTTGTGTAAGATTAATTCATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1435 GAGTTTGTGTAAGATTAATTCATGATGATGATGATGATGATGATGATGATGATGAT 1494
QY 1381 TTGTTGAAACCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1495 TTGTTGAAACCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1554

```

QY 1441 AAATTTAACTCGGGGACTACTATGATGATCTATGTTTGAAGTACTTGAAGAAGTG 1500
DB 1555 AAATTTAACTCGGGGACTACTATGATGATCTATGTTTGAAGTACTTGAAGAAGTG 1614
QY 1501 GAGGTAGTCAAGGTTCTCTTTAGTCTGCTGCAACTATGGCAAGATTGGACCGAG 1560
DB 1615 GAGGTAGTCAAGGTTCTCTTTAGTCTGCTGCAACTATGGCAAGATTGGACCGAG 1674
QY 1561 CATGTGAAGCTGATGCTATGACAGGCACTGAGAAAGTTTCTCTCCGTATACAGAT 1620
DB 1675 CATGTGAAGCTGATGCTATGACAGGCACTGAGAAAGTTTCTCTCCGTATACAGAT 1734
QY 1621 AGAACTCGGCTGACCCCAAGATGTCAGAGACAGTGTATGATGATCTGTTGT 1680
DB 1735 AGAACTCGGCTGACCCCAAGATGTCAGAGACAGTGTATGATGATCTGTTGT 1794
QY 1681 AACATGTAGGCGCTGATGCTGAGCTGCTGTTTATGCAAGAGCTGTAAGCCTCT 1740
DB 1795 AACATGTAGGCGCTGATGCTGAGCTGCTGTTTATGCAAGAGCTGTAAGCCTCT 1854
QY 1741 GAAACTTTGAACTATGATATGATATGCAATTGAGCTGGGCTTCAGAGAGTGTGAT 1800
DB 1855 GAAACTTTGAACTATGATATGATATGCAATTGAGCTGGGCTTCAGAGAGTGTGAT 1914
QY 1801 GAAACTACTGTGAAATGTCGCTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA 1860
DB 1915 GAAACTACTGTGAAATGTCGCTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA 1974
QY 1861 GCTGCTGCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1975 GCTGCTGCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2034
QY 1921 AGCTCATCTTTTAAAGCAAGAGATATGCTTCTTATGCAATGATGATGATGATGAT 1980
DB 2035 AGCTCATCTTTTAAAGCAAGAGATATGCTTCTTATGCAATGATGATGATGATGAT 2094
QY 1981 GGGTCACTGAGCTGCAAGTGAAGCACTTCCCAAGATGATGATGATGATGATGATGAT 2040
DB 2095 GGGTCACTGAGCTGCAAGTGAAGCACTTCCCAAGATGATGATGATGATGATGATGAT 2154
QY 2041 AATATAGTATCCAAAGTGGCAAGATTAAGTCTTGGCTTTGGGCTGATCAACCGATA 2100
DB 2155 AATATAGTATCCAAAGTGGCAAGATTAAGTCTTGGCTTTGGGCTGATCAACCGATA 2214
QY 2101 GAAATGTATCCAAAGTGGTGGATGAGGCGCAATGCTGAAGATTTGCACTGACAGCAGCT 2160
DB 2215 GAAATGTATCCAAAGTGGTGGATGAGGCGCAATGCTGAAGATTTGCACTGACAGCAGCT 2274
QY 2161 GAAACTGCGCAGCTGGGTTGTTATGATATTAACACTGTTGAAACTATGTTGACAGT 2220
DB 2275 GAAACTGCGCAGCTGGGTTGTTATGATATTAACACTGTTGAAACTATGTTGACAGT 2334
QY 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGGTGAAGCAACTGAGAGAGTCTCT 2280
DB 2335 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGGTGAAGCAACTGAGAGAGTCTCT 2394
QY 2281 TGTCTATCTGATTTGGTCTTCAACGAAACAAATGCTATGATGCAAGACTTACACACA 2340
DB 2395 TGTCTATCTGATTTGGTCTTCAACGAAACAAATGCTATGATGCAAGACTTACACACA 2454
QY 2341 AGATAGGAAGTTTCTGCTGCAAGTCAAGGTTGAAAACTGTAAGGCTGTGTTCTTGA 2400
DB 2455 AGATAGGAAGTTTCTGCTGCAAGTCAAGGTTGAAAACTGTAAGGCTGTGTTCTTGA 2514
QY 2401 TCATAA 2406
DB 2515 TCATAA 2520

```

RESULT 6
ADJ38130
ID ADJ38130 standard; DNA; 3667 BP.
XX

```

AC ADJ38130;
XX
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana Arc6-1 genomic DNA Segid2.
XX
KW prokaryotic type; plasmid division; Fun2; ARC6; ARCS; Pzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; ds.
XX
OS Arabidopsis thaliana.
XX
PN W02004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003MO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Oosteryoung KM, Vitla S, Koksharova OA, Gao H;
PI WPI, 2004-082486/08.
XX
DR P-PSDB; ADJ38202.
XX
PT New isolated Fun2, ARCS and/or Pzo-like nucleic acid sequences, useful
PT for further characterizing plasmid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 2; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fun2 (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.
XX
SQ Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;
XX
Query Match 71.7%; Score 1724; DB 12; Length 3667;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 1 ATGGAAGCTCTGATGATCAGTCCGATGCTCTCTCCCATTCATATGACGATTAACA 60
DB 481 ATGGAAGCTCTGATGATCAGTCCGATGCTCTCTCCCATTCATATGACGATTAACA 540
QY 61 CCGGCGACGCAAAAGCTCCGAGTACGACCAACACCTCTTACACTATCTCTCCGACG 120
DB 541 CCGGCGACGCAAAAGCTCCGAGTACGACCAACACCTCTTACACTATCTCTCCGACG 600
QY 121 AATGAGGCGGACCGCTCTCTCTCGGACTTCAATTTGACCTCGGATTCCTCTCTCTCC 180
DB 601 AATGAGGCGGACCGCTCTCTCTCGGACTTCAATTTGACCTCGGATTCCTCTCTCTCC 660
QY 181 TTGCGCACGCGCACACACACCGGCACTCTGCTCTCTGCGACACATCTATGATGCTCC 240
DB 661 TTGCGCACGCGCACACACACCGGCACTCTGCTCTCTGCGACACATCTATGATGCTCC 720
QY 241 GAAAGCGAGTCTCCCATTCCTCCATTTATTTCTTACAGAGTATGAGAGCTCAAAACATTTTC 300
DB 721 GAAAGCGAGTCTCCCATTCCTCCATTTATTTCTTACAGAGTATGAGAGCTCAAAACATTTTC 780
QY 301 TTAAACGATGATGATGAAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 360

```

Db 781 TTBACGATGGAATGAGAGACATTCGAGCTAGGGTTTCGAAACCGCGCAATTCGGT 840
Qy 361 TTCAGGAGAGAGCGCTTTAATCAGCCGAGAGACATTCCTCAAGCTGCTGGCAACTC 420
Db 841 TTCAGGAGAGAGCGCTTTAATCAGCCGAGAGACATTCCTCAAGCTGCTGGCAACTC 900
Qy 421 TCTAATCCTCGGCTGAGAGAGATCAATGAAGCTCTTCTGATGATGAAGAGCTACA 480
Db 901 TCTAATCCTCGGCTGAGAGAGATCAATGAAGCTCTTCTGATGATGAAGAGCTACA 960
Qy 481 GTTCATCATGATGCTCTGGGAT-----AAGGTTCTGGGC 504
Db 961 GTTCATCATGATGCTCTGGGATGAAGTATTCGATTCGGAATATAAGTTTCTTC 1020
Qy 505 -----AAGGTTCTGGGC 518
Db 1021 GTTTAATTTTCATGAATGGATTAAGAAAGAACTTTATCTAGTGAGAGTTCTGGGC 1080
Qy 519 TCTCTGTATTTGCAAGAAAGGTGTGAGACTGATAGTTCTCGGGTTGGTAGGCTCT 578
Db 1081 TCTCTGTATTTGCAAGAAAGGTGTGAGACTGATAGTTCTTCGGGTTGGTAGGCTCT 1140
Qy 579 GCTTAAGAGAGGTTGCTTAAGCTGTTTAAAGCAAGATGTGTTTAAATGAGCGCTGC 638
Db 1141 GCTTAAGAGAGGTTGCTTAAGCTGTTTAAAGCAAGATGTGTTTAAATGAGCGCTGC 1200
Qy 639 GTTTCGATGCTCGAGGAGTCTATGAGCTTGAATCCACTGATTTTATCTGTTA 698
Db 1201 GTTTCGATGCTCGAGGAGTCTATGAGCTTGAATCCACTGATTTTATCTGTTA 1260
Qy 699 TGAGTTGTTGAGAGCTTTGAAGCTTTTAC----- 720
Db 1261 TGAGTTGTTGAGAGCTTTGAAGCTTTTACAGAGTATGACTTGTGTAATTG 1320
Qy 731 -----A 731
Db 1321 ACAGAGGTTGGCTTTAAGAACTTCTGATTTGATTTGATTTGATGCTGTGTA 1380
Qy 732 GAGAGAGAGAGAGATGACCTTGACCCGATTTACGTGCAAAATGATGAGCTTTGA 791
Db 1381 GAGAGAGAGAGAGATGACCTTGACCCGATTTACGTGCAAAATGATGAGCTTTGA 1440
Qy 792 AAGATCACTCCCGCTTATGCTGAGCTACTTGGCTTACCGCTTGGTATGATTAAGC 851
Db 1441 AAGATCACTCCCGCTTATGCTGAGCTACTTGGCTTACCGCTTGGTATGATTAAGC 1500
Qy 852 TCGAAAAAGACTAATGCTTTAAGCGGTGCGGAATATTTGTCTGTGAGAGAG 911
Db 1501 TCGAAAAAGACTAATGCTTTAAGCGGTGCGGAATATTTGTCTGTGAGAGAG 1560
Qy 912 TGAGCATCAGCTCTGTTGGGGTTTGAACCGGTGAGAGATTTATGATGAGCGTTT 971
Db 1561 TGAGCATCAGCTCTGTTGGGGTTTGAACCGGTGAGAGATTTATGATGAGCGTTT 1620
Qy 972 ACGAATGACGCTGCTGAGC----- 991
Db 1621 ACGAATGACGCTGCTGAGCATTAAGATTAATCTTTTAAATTTCTTTAGCA 1680
Qy 992 -----AGTTGATCTTTTGA 1008
Db 1681 TGAATTAATCTTGAAGTTTCTCATTTTAATGATGCTGTGAGATTTGATCTTTTGA 1740
Qy 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGCACTTGTCT 1068
Db 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGCACTTGTCT 1800
Qy 1069 GTGGCTCAAGCTTTTATTTGTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1128
Db 1801 GTGGCTCAAGCTTTTATTTGTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1860
Qy 1129 CAGCACTTCAAGAGGCTAAGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 1188

Db 1861 CAGCACTTCAAGAGGCTAAGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 1920
Qy 1189 CGAAATTAATTTGGAGATTAAGCTTCCGTACAGAAAGGAGACTGTGTCATCTGTAATAGC 1248
Db 1921 CGAAATTAATTTGGAGATTAAGCTTCCGTACAGAAAGGAGACTGTGTCATCTGTAATAGC 1980
Qy 1249 AAAAGTGAATATGCCGTATGTGTTGGCTTACAGAGTGAAGATTTACAAATATAGAAAT 1308
Db 1981 AAAAGTGAATATGCCGTATGTGTTGGCTTACAGAGTGAAGATTTACAAATATAGAAAT 2040
Qy 1309 CCAAGTATGTGAGATTTGTTTGAAGATTAATCCGATGACAAATGATGATCTCCCT 1368
Db 2041 CCAAGTATGTGAGATTTGTTTGAAGATTAATCCGATGACAAATGATGATCTCCCT 2100
Qy 1369 GGAATATGCAAAATTTGTTGAAAACCTGTTGGAGGGGTTGTCTTCTAGTTCAAGAGC 1428
Db 2101 GGAATATGCAAAATTTGTTGAAAACCTGTTGGAGGGGTTGTCTTCTAGTTCAAGAGC 2160
Qy 1429 ACCAAAGATTAATAATTTAACTCGGGACATCATGATGATCTATGTTTGAATTAC 1488
Db 2161 ACCAAAGATTAATAATTTAACTCGGGACATCATGATGATCTATGTTTGAATTAC 2220
Qy 1489 TTGAAAAGTGAAGTACTTCAAGGTTCTCCTTTAGCTGCTGCAACTATGGAAG 1548
Db 2221 TTGAAAAGTGAAGTACTTCAAGGTTCTCCTTTAGCTGCTGCAACTATGGAAG 2280
Qy 1549 ATTGAGCCGAGAGATGAAAAGCTGAGTATGACAGGCACTGCAAAATTTTCTCTCC 1608
Db 2281 ATTGAGCCGAGAGATGAAAAGCTGAGTATGACAGGCACTGCAAAATTTTCTCTCC 2340
Qy 1609 CGCTATACAGATTAAGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTATTAGTGA 1668
Db 2341 CGCTATACAGATTAAGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTATTAGTGA 2400
Qy 1669 GATCTGTGTGTAACAATGAGCGCTGATGTGAGACCTGTGTTTATGCAAGAGCT 1728
Db 2401 GATCTGTGTGTAACAATGAGCGCTGATGTGAGACCTGTGTTTATGCAAGAGCT 2460
Qy 1729 GTAAGACCCCTGAAAACCTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCAGAG 1788
Db 2461 GTAAGACCCCTGAAAACCTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCAGAG 2520
Qy 1789 AGTAGCGTTGATGAAACTACTGTGTAATGTCCGTTGCTGATATGTTAAAGAGCAAGT 1848
Db 2521 AGTAGCGTTGATGAAACTACTGTGTAATGTCCGTTGCTGATATGTTAAAGAGCAAGT 2580
Qy 1849 GTGAAGATCCTAGCTCTGCTGTGAGCAATTTGACTAATTCATCTGACCCGCAAGTAT 1908
Db 2581 GTGAAGATCCTAGCTCTGCTGTGAGCAATTTGACTAATTCATCTGACCCGCAAGTAT 2640
Qy 1909 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGAT 1968
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGAT 2700
Qy 1969 GTGCTTACCA----- 1978
Db 2701 GTGCTTACCAATGATTAATGATGCAATTTTCATATATCTGATGCTCAAAATA 2760
Qy 1979 ----- 1978
Db 2761 TGCTGTTTGTGAGCTAAGAACATAGTCCACTTAATACATGTCCAAAAGTTGACC 2820
Qy 1979 ----- 1978
Db 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGTGAATTTTGTGATCAAA 2880
Qy 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATACGTAGATAG 2940
Qy 1979 -----T 1979
Db 2941 AGATTGCTTATGATGAGCTTTGTGCAACTTTTCTTCTGATTTTCTTTTCAATTT 3000

QY 1980 AGGTCAGTCAGAGCTGACGACGATTGAGAGCACTTCCAGATGATGCTAGACCTGACGAGA 2039
 DB 3001 AGGTCAGTCAGAGCTGACGACGATTGAGAGCACTTCCAGATGATGCTAGACCTGACGAGA 3060
 QY 2040 GAATATATGATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGGCTGATCAACCGCAT 2099
 DB 3061 GAATATATGATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGGCTGATCAACCGCAT 3120
 QY 2100 AGAATGTTACGAG----- 2113
 DB 3121 AGAATGTTACGAGGTCAGGAAATTAATCTAACATTCATCAATTGTTGAAAACGTG 3180
 QY 2114 -----AGTTTGGAGTGGC 2128
 DB 3181 TGAGACATGATATAGTCTGCTGCTTGTGATTCTGTTATTTATAGGTTTGGATGGGC 3240
 QY 2129 GAATGCTGAAGATTGAGCTGACAGAGCAGCTGAAACTGGCAGCTTGGGTTGTTATG 2188
 DB 3241 GAATGCTGAAGATTGAGCTGACAGAGCAGCTGAAACTGGCAGCTTGGGTTGTTATG 3300
 QY 2189 ATTATACACTGTGAAACTATCTGTGACAGTGTGACAGTCTGACAGATGAAACCGCTG 2248
 DB 3301 ATTATACACTGTGAAACTATCTGTGACAGTGTGACAGTCTGACAGATGAAACCGCTG 3360
 QY 2249 CTCTGCTGAGAGCACTGAGAGAGTCTGCTGATCTGATCTGATCTGATCTGATCTGATCTG 2308
 DB 3361 CTCTGCTGAGAGCACTGAGAGAGTCTGCTGATCTGATCTGATCTGATCTGATCTGATCTG 3420
 QY 2309 ACAATGCTACTGATGTGACAGAACTGACACAAAGATTAAGATTTCTGCTGCAAGTCA 2368
 DB 3421 ACAATGCTACTGATGTGACAGAACTGACACAAAGATTAAGATTTCTGCTGCAAGTCA 3480
 QY 2369 GGTGGAATCACTGAAAGGCTCTGTTCTTGCATCATTA 2406
 DB 3481 GGTGGAATCACTGAAAGGCTCTGTTCTTGCATCATTA 3518

RESULT 7
 ADJ38136
 ID ADJ38136 standard; DNA; 3667 BP.
 AC ADJ38136;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Arabidopsis thaliana AtFen2 genomic DNA SeqID10.
 XX
 KW prokaryotic type; plasmid division; Fen2; ARC6; ARC5; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402342P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Osteoyoung KW, Viltha S, Koksharova OA, Gao H;
 DR WPI, 2004-082486/08.
 DR P-PSDB; ADJ38203.
 XX
 PT New isolated Fen2, ARCS and/or Fzo-like nucleic acid sequences, useful
 for further characterizing plasmid division in plant cells, and in

PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Claim 1; SEQ ID NO 10; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plasmid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fen2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plasmid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a DNA sequence
 CC which is related to the invention.
 XX
 SQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match 71.5%; Score 1720.8; DB 12; Length 3667;
 Best Local Similarity 79.1%; Pred. No. 0;
 Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

QY 1 ATGGAAGCTGTGATGACGCTGCGCATTTGCTCTCCCATTTCAATTATGCGATTACCA 60
 DB 481 ATGGAAGCTGTGATGACGCTGCGCATTTGCTCTCCCATTTCAATTATGCGATTACCA 540
 QY 61 CCGGCGACGACAAAGCTCCGAGCTGACGACCAACACCTCTACACTATCTGCGCGCAGC 120
 DB 541 CCGGCGACGACAAAGCTCCGAGCTGACGACCAACACCTCTACACTATCTGCGCGCAGC 600
 QY 121 AAATGGGCGGACCGCTCTCTCTCCGACTTCAATTGACCTCCGATTCTCTCTCTCC 180
 DB 601 AAATGGGCGGACCGCTCTCTCTCCGACTTCAATTGACCTCCGATTCTCTCTCTCC 660
 QY 181 TTGGCCACCGGACGACGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 DB 661 TTGGCCACCGGACGACGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 QY 241 GAAAGCGACGCTCCGATCCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 300
 DB 721 GAAAGCGACGCTCCGATCCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 780
 QY 301 TTAAACGATGATGATGAAAGAGATTCGAAGCTAGGTTTCCAAACCGCGCAATTCCGT 360
 DB 781 TTAAACGATGATGATGAAAGAGATTCGAAGCTAGGTTTCCAAACCGCGCAATTCCGT 840
 QY 361 TTCAGCGAGAGGCTTTAATCAGCCGAGACAGATTTCTCAAGCTCTTGCAGAACTCTG 420
 DB 841 TTCAGCGAGAGGCTTTAATCAGCCGAGACAGATTTCTCAAGCTCTTGCAGAACTCTG 900
 QY 421 TCTAATCCTCGCTCTGAAAGAGATTCATGAAGGCTTCTGATGATGAAGAGCTACA 480
 DB 901 TCTAATCCTCGCTCTGAAAGAGATTCATGAAGGCTTCTGATGATGAAGAGCTACA 960
 QY 481 GTCATCACTGATGCTCTTGGGAT----- 504
 DB 961 GTCATCACTGATGCTCTTGGGATTAAGTAATTTCGATTTCCGAATTAATAAGTTCTTC 1020
 QY 505 -----AAGTTCTGCGGC 518
 DB 1021 GTTTTAATTTCAATGATGATGATTAAGAAAGAACTTTTATCTAGTGAAGGTTCTGCGGC 1080
 QY 519 TCTCTGTGATTTGCAAGAGGCTGAGACTGAGATGATTTCTGCGGTTGAGAGCTCTC 578
 DB 1081 TCTCTGTGATTTGCAAGAGGCTGAGACTGAGATGATTTCTGCGGTTGAGAGCTCTC 1140
 QY 579 GCTTAAGAGAGGTTGCTTAAGCTTTTAAAGCAAGATGATTTTAAAGTGGCGCTTGC 638
 DB 1141 GCTTAAGAGAGGTTGCTTAAGCTTTTAAAGCAAGATGATTTTAAAGTGGCGCTTGC 1200
 QY 639 GTTTCGATGCTCTGAGGAGATGATGAGCATTTGATTCACCTGATTTTAACTGTTTA 698
 DB 1201 GTTTCGATGCTCTGAGGAGATGATGAGCATTTGATTCACCTGATTTTAACTGTTTA 1260

Db 3421 ACAATGCTACTGATGTCAGAACCTACACAGATACGAGTTTCTGTGCTCCAGTCAG 3480
QY 2369 GGTGGAATCACTGAGAGCTCTGTTCTTCATCATTA 2406
Db 3481 GGTGGAATCACTGAGAGCTCTGTTCTTCATCATTA 3518

RESULT 8
ADJ38212/c
ID ADJ38212 standard; cDNA; 561 BP.
XX
AC ADJ38212;
XX
DT 06-MAY-2004 (first entry)
DE Plastid division-related Arc6 orthologue cDNA 5.
XX
KM prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003MO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oesteryoung KM, Viltha S, Koksharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
PT New isolated Fun2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fun2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 561 BP; 167 A; 136 C; 107 G; 148 T; 0 U; 3 Other;

Query Match 20.0%; Score 482; DB 12; Length 561;
Best local Similarity 99.4%; Pred. No. 1.5e-130;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1922 GCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATGATGCTGCTACCATAG 1981
Db 561 GCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATGATGCTGCTACCATAG 502

QY 1982 GGTGAGTCAAGCTGAGATGATTCAGAGCACTTCCGAAATGATGCTGAGAGCTGCGAGA 2041
Db 501 GGTGAGTCAAGCTGAGATGATTCAGAGCACTTCCGAAATGATGCTGAGAGCTGCGAGA 442

QY 2042 ATATGATATCAAGTGGCAGAAATTAAGTCTGCTGTTGGGCTGATCAGCGCATAG 2101
Db 441 ATATGATATCAAGTGGCAGAAATTAAGTCTGCTGTTGGGCTGATCAGCGCATAG 382

QY 2102 AAATGTTACAGAGGTTTGGATGGGCGAATGCTGAAGATTTGGAATGACAGAGCAGCTG 2161
Db 381 AAATGTTACAGAGGTTTGGATGGGCGAATGCTGAAGATTTGGAATGACAGAGCAGCTG 322

QY 2162 AAATGCGAGCTTGGGTTGGTTTATGATTAATCACTGTTGAATATCTGTTGACAGTG 2221
Db 321 AAATGCGAGCTTGGGTTGGTTTATGATTAATCACTGTTGAATATCTGTTGACAGTG 262

QY 2222 TGAAGCTTCAGAGATGGAACCCGCTGCTGTTGGAAGCAATCTGGAAGAGTCTGCTT 2281
Db 261 TGAAGCTTCAGAGATGGAACCCGCTGCTGTTGGAAGCAATCTGGAAGAGTCTGCTT 202

QY 2282 GTCTATCTGATTTGGTTTCATCCGAAACAAATGCTCTGATGTCAGAACTTACACAA 2341
Db 201 GTCTATCTGATTTGGTTTCATCCGAAACAAATGCTCTGATGTCAGAACTTACACAA 142

QY 2342 GATACGAAGTTTCTGCTCCAACTCAGGCTGGAATAATCACTGAAGCTCTGTTCTTGCAT 2401
Db 141 GATACGAAGTTTCTGCTCCAACTCAGGCTGGAATAATCACTGAAGCTCTGTTCTTGCAT 82

QY 2402 CATTA 2406
Db 81 CATTA 77

RESULT 9
ADJ38206
ID ADJ38206 standard; DNA; 2283 BP.
XX
AC ADJ38206;
XX
DT 06-MAY-2004 (first entry)
DE Plastid division-related Arc6 orthologue gene 1.
XX
KM prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ds.
XX
OS Oryza sativa.
XX
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003MO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oesteryoung KM, Viltha S, Koksharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
DR P-PSDB; ADJ38205.
XX
PT New isolated Fun2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fun2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used

as herbicide targets. The present sequence is that of a gene which is related to the invention.

Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 20.0%; Score 481; DB 12; Length 2283;

Best Local Similarity 54.3%; Pred. No. 6,6e-130;

Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
QY 108 CTGCTCCGCGAAGAAATGGGCGGACCGCTCTTCTCTCGAATTGCAATTTCACCTCCGATTC 167
DB 132 CTGGGCGGAAGCGCTCTTGGCGGACCTTCCACCTCCACCGCGGCTCCGACCC 191
QY 168 CTGCTCTCTCTCTCTGCGACCGGACGACGACGACGACGACGACGACGACGACGACGACGAC 227
DB 192 GCGCTCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 251
QY 228 TATTGATCGTCCGGAAGCGGACGCTCCCATCCCATTTGATTCTTACAGGATTTAGAGAGC 287
DB 252 CCGGAGCGCGGGAAGCGCTCCCTCCGCTCCAGTGCATTTCTACAAAGGTTCTAGGGGCG 311
QY 288 TCAACACATTTCTTAAACGATGGAATCAGAGAGGATTTGAAAGTGGTTTCGAAGC 347
DB 312 AGAGCCACATTTCTTGGCGATGCGATCAGAGAGGCGTTGAGGCGACGATAGCAGAGCC 371
QY 348 GCGGCAATTCGTTTCAGGACGACGCTTAACTCAGCGGACGACGATTCCTCAAGCTGC 407
DB 372 ACCGCGATATGCTTACGACGAGATGCTCTTGTGCTGCGACAAATGCTGCGAGATGC 431
QY 408 TTGCGAACTCTGTCTAATCTCGGTCTAGAGAGATGACATGAAAGTCTTCTGATGA 467
DB 432 CCATGACATCTCTATGAAACGAAACTCCGCGACTCAGATGATGCTGCTTCTGAGAA 491
QY 468 TGAAGAGCTACAGTCTACTGATGTTCTTGGGATTAAGTTCCTGGGGCTCTGTGT 527
DB 492 CCGTGAAGAGCTCTCACTGATATTTGCTTGGGAAAGAGAGCGCTGG----- 540
QY 528 ATTGCAAGAGGTGAGACTGAGATGTTCTCGGGTTGGTGAAGGCTGCTGCTTAAAGA 587
DB 541 -----GAGGCACTTGCTGTGCTTGTGTAACCTGAGAAACAGTTGCTTCTGGA 584
QY 588 GAGGTGCTTAAGTCGTTTAAGCAAGATGTGTTTAAGTATGCGCTGCTGTTTCTGCA 647
DB 585 TCGGCGACCGAAGCGCTTCAAGGACGAGCGTGTGCTAGCGATGGCTCTGCTTATGAGGA 644
QY 648 TGTCTGAGGAGTGTATGCGATTTGATCCACTGATTTTATTACTGTTATGAGTTGT 707
DB 645 TCTATCAAGGAGTGTATGCGACGACGACCTCCAGATGTAATGGCTGCGAGGTGCT 704
QY 708 TGAAGAACTTTGAAGCTTTTACAGAGGAAAGGAGAAAGTACCTTGCACCGGATTTAGC 767
DB 705 CGAGAGGCTCTCAAGCTCTTGCAGAGAAATGAGCAAGCAATCTCCACCTGATCTGCT 764
QY 768 TGACAAATTTGATGAGACTTTGGAAGAGATCACTCCGCTTATGTTCTTGAAGCTACTTGG 827
DB 765 TTACAGATTTGAGAACTCTCAGAGAGATTACCTCGCTGTATTTGAGAGCTTCTCTC 824
QY 828 CTTACCGCTTGTGTGATGATACGCTGCGAAAAGCTAATGTGTTTAAAGCGTGTGCGAA 887
DB 825 CTTTCAATTTGACACAGAGATCATTAAGAGCGCAAGAGGCGTTCAAGGTGCGAAGAA 884
QY 888 TATTTTGTGTCTGTGAGAGAGGTGAGCATGAGCTCTTGTGGGGGTTTGAACCGCTGA 947
DB 885 CATTTTGTGAGCGTTTGGAGAGAGGTATTTGCTACCGTTGAGAGAGATTTTCTCTGGA 944
QY 948 GAAGTTATGATGAGCGTTTATTAAGATGACAGCTGCTGAGAGGTTGATCTTTTGT 1007
DB 945 AGCTTATGATGAGCGCTTTTGAAGATGACATCAATTTGAGAGATGATTTCTTTC 1004
QY 1008 AGCTACCCCAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCT 1067
DB 1005 AAAAAGCCGAAATGCAATTCCTCTGATGTGTTGAATTTTACAAATGATGACATTTGACA 1064
```

```
QY 1068 TGTGCTCAAGCTTTTATTTGTTAGAGGACCACTTTTACAGAGATGCTGATTAAGCAATT 1127
DB 1065 TGTGCTCAAGCAATTTATTAAGGCGCAATTTCAATCATGATGGCGAGATCTTTT 1124
QY 1128 CCAGCACTTTCAGAGCTTAAGGTATGAGCTATGAGATTCCTGCGATGTGTATGATAC 1187
DB 1125 TGAACACTTCAGAAATTCA-----CATAGTTTCTATTA 1160
QY 1188 ACGGAATTAATTTGGAGATGAGCTTGGCTAGAGAGGAGGAGCTGTGACCTGTTTAAAG 1247
DB 1161 TGCTTATGATTAATGAGATGAGACTTTCATTTGAAAGGAGATCTGCTCATTTGCTATCGG 1220
QY 1248 CAAGTTGATGAATGCGGATGTGTTGGGCTTACAGACTGAGAGATTCACATATAGAA 1307
DB 1221 AGATTTTACCAAGTGCAGATGTGGCTTGAAATGATTAATGATCTTCAACATACAGAGA 1280
QY 1308 TCCAGCTATTTGAGATTTGTTTGAAGAAATTCAAATC---GTATGACAAATGATGATCT 1364
DB 1281 CCGCAAAATTTAGAAATTTATGTTGACCAACTTCACATCAATGAAAGAGATGATCTTCT 1340
QY 1365 CCTGACTATGCAAAATTTGTGAAACCTGTTGCGAGGGGTTGCTTCTTCTAGTTTCAG 1424
DB 1341 TCCAGGGCTGTGCAACTTTTGAAGCTTGGCTTATCTTTGAGGTTTCTTCTAGAGCAG 1400
QY 1425 AGACACCAAGATAAAAATTTAACTCGGGAATCTATATGATGATCTATGTTTGAAG 1484
DB 1401 AGATATCTGGGGGCAATGCACTTCACTGAGATTTCTAGATGATCAGAAAGTTTAAAG 1460
QY 1485 TTTCTTGAAGAGTGAAGTATGTTCAAGGTTCTCTTTAGCTGCTGCTGCAACTATGAG 1544
DB 1461 CTACCTTGAAGAGATGAGAGGTGTGTCTTCTCAATTTGGCTGCTGCTGCTATATGC 1520
QY 1545 AAGGATTTGAGCGCGACATGTGAAAGCTATGCTATGACAGGACTGCAAGAAATTTTTC 1604
DB 1521 AAACCTTGTGCTCAAGCTTACAGCTGCACTTGGTA----- 1555
QY 1605 TTCCGCTATACAGATGAAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTTCAG 1664
DB 1556 -----CTGTGAATTCAAATGCTATTTCAAGCG----- 1581
QY 1665 TGTAGATCTGTGTTGTTAACAATGTAGCCGTGATGAGTGTGAGCCCTGTTTATTTGACA 1724
DB 1582 -----TTCAACAA 1589
QY 1725 AGCTTAAGACCTCTGAAACCTTTGAAACTAATGATTAATGCAATTGAGCTGGGCTCTC 1784
DB 1590 GCTTTTTCATTAATGAAACAGTTTGAACAGTCAAGCAGCAAGAAATTAATAAGATGGCC 1649
QY 1785 AGAGATGAGCTTGAATGAATCTGTTGAATGTCGTTGCTGATGATGTTAAAGAGGC 1844
DB 1650 TGGGGGATATCTTGAATAATTTTGAACAGAAATGCACTGCTCATGATTCGAATATGC 1709
QY 1845 AAGTGAAGATCTTACGCTGTGTGTGCAATTGACATGATTTCACTGTTCAAGCAAG 1904
DB 1710 CGCTTGAAGATATATCTGCTGCGCACTGTTTGAACGTTGGCAGTAAATGGGCGCAA 1769
QY 1905 GTATTTTCTTAAAGAGCTCATCTTTCAAGCAGAGATGATGTTCTTCTTATGAATC 1964
DB 1770 ATATTTGCTCTGTAAGAGGCGCTTCTGTCTATTAAGAGTGAATGATCTGTGGCAGT 1829
QY 1965 TGAT-----GTGCTACATTAAGGCTCAGTCAAGCTGAGATTCAGAAAGCACTTCC 2015
DB 1830 TGCTTAATGATGATCTTATGATGATCTGCTGCAATGATGAAGATCCAGTATCATATTC 1889
QY 2016 CAGATGATGCTTGAAGCTGCGAGAAATTAATGATCCAAAGTGCAGAAAGTAAATGCTCT 2075
DB 1890 TAGAATGATGCGAACTGCGAGAAATATTTGTTCCCAAGTGCAGAGATTAATTAATCT 1949
QY 2076 GGTCTTTGGGCTTCAACCCGATAGAAATGTAACAGAGGTTTGGATGGGGAAGTCT 2135
DB 1950 GGGCTTTGGGACCAAGCAATTCGTTGCTATTCATTTGCAAGAGGTTCTTGAATGGCAATGCT 2009
QY 2136 GAGATTTGAGCTGACAGAGCACTGAAACTGCGCAGCTTGGGTTGTTATGATTAATAC 2195
```

```

Dc 2010 AAAGTGTGACCTGACGAGAGCGAGATTGACGCTGATGGGTGTTCTGGAGATATAC 2069
Qy 2196 ACAGTTGAACCTATCTTTGACAGTGTGACAGTCTGACAGANTGAAACCGTGTCTGT 2255
Db 2070 ACTATCCGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2129
Qy 2256 GGAAGCACTCTGAGAGAGTCTGTTCTATCTGATTTGGTTCATCCAGAAACAATGC 2315
Db 2130 GAGGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2189
Qy 2316 TACTGATGTGAGAACTTACACACAGATGAGAAAGTTTCTGTCGAAG--TCAGGGTG 2372
Db 2190 TTGATATGACACAAATACACTACCTCCGATGATGATGATGATGATGATGATGATGATGAT 2249
Qy 2373 GAAATCATGTAAGGCTCTGTTCTTGTGATCATTA 2405
Db 2250 GAAATTAACGAGAGAGAGAGCTCTCAAGTGTGA 2282

RESULT 10
ADJ38264
ID ADJ38264 standard; cDNA; 631 BP.
XX AC ADJ38264;
XX DT 06-MAY-2004 (first entry)
XX DE Plastid division-related Arc6 orthologue cDNA 51.
XX KM prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX KM herbicide target; gene; ss.
XX OS Prunus persica.
XX PN WO2004001003-A2.
XX DT 31-DEC-2003.
XX PF 20-JUN-2003; 2003WO-US019536.
XX PR 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-0060070.
XX PA (UNMS ) UNIV MICHIGAN STATE.
XX PI Osteryoung KM, Vicha S, Koksharova OA, Gao H;
XX DR WPI; 2004-082486/08.
XX PT New isolated Fun2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plastid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX PS Disclosure; Fig 8; 287bp; English.
XX OS This invention relates to novel prokaryotic type or plastid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Fun2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plastid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is a cDNA sequence which is
XX CC related to the invention.
XX SQ Sequence 631 BP; 174 A; 123 C; 155 G; 178 T; 0 U; 1 Other;
XX Query Match 13.7%; Score 328.8; DB 12; Length 631;
XX Best Local Similarity 71.5%; Pred. No. 1.8e-85;

```

```

Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;
Qy 913 GAGACATCAGCTCTTGTGGGGGTTTGACCCGTGAGAACTTTATGAAATGAGGCGTTT 972
Db 1 GCAGTGCATTTGTGGGGGNGATTCTACGTGAGAAATTTCAATGAAAGGCGTTT 60
Qy 973 CGAATGACAGCTGTGAGAGGTTGATCTTTTGTAGCTACCCCAAGCAATATTCAGCA 1032
Db 61 CATATGACCTGAGTGAGAGGTTGATTTATTTGTAGCTACCCCAATATATCCCGCA 120
Qy 1033 GAGTCATTTGAAGTTTACGAAAGTTGACCTGCTCTTGTGGCTCAAGCTTTATTTGTTAG 1092
Db 121 GAAAGCTTTGAAGTTTATGAGGGTGGCTCTTCCCTGTTCTCAAGCTTTGTTGTTAA 180
Qy 1093 AAGCACACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAGAGGCTAAGTA 1152
Db 181 AAACCTCATACATTTCAAGATGCTGAAACCTTATTCAGAACTTCAGAGCTTAAGTA 240
Qy 1153 ATGCTATGAGAGATTCCTGCGATGTTGTATGATACACGAAATTAATGGAGATAGACTTC 1212
Db 241 ACAGCTGTAGACATTTCTTTGACAACTATATTAACCAAGAAAGAGAGATAGACTTT 300
Qy 1213 GGTCTAGAAAGGAGACTCTGTGACCTGCTTATAGGCAAGTTGATGATGATGATGATGATG 1272
Db 301 GCTTTGAGAGGGGAGACTGTGTCATCTTCTAGGGGACCTGATGACAGTCTGCTGG 360
Qy 1273 TTGGGCTTACAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1332
Db 361 TTGGGCTTACAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Qy 1333 GAGAACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 421 GAGAACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 1381 TTGTTGAAACCTGTTGGAGGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db 481 CTATGGAAGAGCTGTGATGAGGGTGTATTCCTCCAGGTTTATGAGACCAAGACATTA 540
Qy 1441 AAATTTAACTCGGGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 541 GAGTTGAGACTGGAGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 1501 GAGTAGTTCAAGGTTCTCCTTAGCTGCTG 1531
Db 601 GATGGCACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 631

RESULT 11
ADJ38215
ID ADJ38215 standard; cDNA; 660 BP.
XX AC ADJ38215;
XX DT 06-MAY-2004 (first entry)
XX DE Plastid division-related Arc6 orthologue cDNA 8.
XX KM prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX KM herbicide target; gene; ss.
XX OS Medicago truncatula.
XX PN WO2004001003-A2.
XX DT 31-DEC-2003.
XX PF 20-JUN-2003; 2003WO-US019536.
XX PR 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-0060070.
XX

```

PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 XX
 PT New isolated Fuz2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fuz2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.
 XX
 SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
 XX
 Query Match 11.6%; Score 278.8; DB 12; Length 660;
 Best Local Similarity 67.3%; Pred. No. 9.5e-71;
 Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;
 XX
 QY 81 ACCTACCCACACACCTCTCACTATCTGCTCCGACGAAATGGGCGACGCTTCT 140
 DB 17 ACCTAACCGTCTCATCTCCCGCCGCTCCGACACGTAATGGGCGAGGACTCAT 76
 QY 141 CTCGCACTTCAATTTCACTCTCGAATCTCTCTCTCTCTCTCTGCGCACCGCCACACAC 200
 DB 77 TTCGATTTTCAATCTCCGCGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
 QY 201 CGGCACTCTGCT 260
 DB 137 AGTCACTCTC-----ACTCTCTCTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
 QY 261 CATTGATTTTCTACAGGATTTAGAGCTCAACACATTTCTTAACCGATGAAATCAGAG 320
 DB 191 TCTCGACCTGTACAAATCTCTGCGCGGCGGAAAGCATTTCTCGGTATGATTTGGAG 250
 QY 321 AGCATTCGAAAGTAAAGGTTTTCGAAACCGCGCAATTTGGTTTTCAGCAGACGCTTAAT 380
 DB 251 AGCTTATGAAGCAAAATTCGAAAGCTCTCTCAATGCTTTCAGTAATGAAGCTTTGAT 310
 QY 381 CAGCCGAGACAGATTTCTCAAGCTCTTGGAAACTCTGTAACTCTGCTCTAGAG 440
 DB 311 TATCTGCTGTCAATTTCTCAAGCTCTTGGAAACCTTACGATCTCTCTCTCTAGAG 370
 QY 441 AGAGTCAATGAGAGTCTTCTGTATGATGAA-----GAGCTACAGTCATCAC 488
 DB 371 AGAGTATATCAAAAGCTCTGTGACGATGAAAGAGATGAGAAATCTTCATTTCTCAC 430
 QY 489 TGATGTTCTTGGAGTAAGGTTCTGTGGGCTCTCTGTGTATTTGCAAGAGGTGTGAG 548
 DB 431 TGAATCTCTTTCGACAAAGTTCTGTGAGCTCTGTGCTGTGTCAGAAAGCTGTGAGAG 490
 QY 549 TGAATGATGTTCTTGGGTTGTGAGGCTCTGTCTTAAGAGAGGTTGCTAAGTCTTAA 608
 DB 491 GAGAGTGTGCTCTTTCGATTTGAGAGGAGGTTTTCAGAGAGGTTTACGAAAGATGTTAA 550
 QY 609 GCAAGATGTGTTTATGATTTAGTGGCGCTTGGCTTCTCGATGTCGAGGAGTGTATGCG 668
 DB 551 GCAAGATGTGTTTATGATTTAGTGGCGCTTGGCTTCTCGATGTCGAGGAGTGTATGCG 610
 QY 669 ATTGATGCACTGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTT 710
 DB 611 TTTGTCT 652

RESULT 12
 ADJ38223
 ID ADJ38223 strand; cDNA; 537 BP.
 XX
 AC ADJ38223;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plastid division-related Arc6 orthologue cDNA 16.
 XX
 KW prokaryotic type; plastid division; Fuz2; ARC6; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; gene; ss.
 XX
 OS Triticum aestivum.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 XX
 PT New isolated Fuz2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fuz2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.
 XX
 SQ Sequence 537 BP; 133 A; 113 C; 151 G; 140 T; 0 U; 0 Other;
 XX
 Query Match 9.3%; Score 224; DB 12; Length 537;
 Best Local Similarity 63.6%; Pred. No. 1.1e-54;
 Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
 XX
 QY 555 AGTTCTTGGGTTGGTGAAGGCTCTGCTTAAGAGAGAGGTTGCTAAGTCTTAAAGCAGA 614
 DB 2 AGTCTTTCGATTTGAGAGGAGGACATTTCTGAGAGACCGCCGCCAAGCGTTCAAGCAGA 61
 QY 615 TGTGTTTATGATTAAGCGCTTGTCTGATGATGTCGAGGAGATGATGCAATTGGA 674
 DB 62 TGTGTTTGTGCAATGAGCGCTGCTTATGATGATGATCAAGGAGACGCAATGCGCGTAG 121
 QY 675 TCACTGATTTTATTAATGCTTATGATGATGATGATGATGATGATGATGATGATGAT 734
 DB 122 CCTTCAGATGATATCGCTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 181
 QY 735 GGAAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
 DB 182 GATGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
 QY 795 GATCACTCCCGGTTATGCTTGGAGCTTATGATGATGATGATGATGATGATGATGATGAT 854

Db 242 TATCACACCTGTTGTTTGGAGCTTTCCTCCCTTGATGAAAAAATCATGAA 301
 Qy 855 GAAAAAGCTAAATGTTTAAACGGTGTGCGCAATATTTTGTCTGTGGAGAGTGG 914
 Db 302 TGAACACCAAGAGGCTTCTGTGTGAGAAACATTTTGTGTGGAGTGTGCGAGAGG 361
 Qy 915 AGCATGAGCTCTTGTGGGGGTTTGAACCGTGAAGATTATGAATAGGCGTTTACG 974
 Db 362 TATTGTAAGTCTTGTGAGAGAGATTTTGCCTGAAGCTTCAATGAATAGCTTCTGCA 421
 Qy 975 AATGACAGCTGTGAGAGAGTGTGATCTTTTGTAGTACCCCAAGCAATATTCAGAGAG 1034
 Db 422 GATGACATGCGCGAGAGATGATTTCTTCTCAAAAACCCGAATAGCATACCGCTGA 481
 Qy 1035 GTCAATTGAAGTTTACCAAGTTGATGCTCTTGTGGCTCAAGCTTTATGTGTA 1090
 Db 482 ATGCTTGAATCTATAGCGTGTGCACTTGCAAAATGTGTCTCAAGCATTTGAAGTA 537

RESULT 13

ACN48855/c
 ID ACN48855 standard; cDNA; 552 BP.
 ACN48855;
 AC 02-DEC-2004 (first entry)
 XX
 DE Cotton primed seed EST clone ID: LIB3825-027-06-N6-H1, SEQ:3636.
 XX
 KM Cotton; EST: expressed sequence tag; transgenic plant; seed;
 KM variety DP50B; library LIB3825; molecular tag; molecular marker;
 KM genetic mapping; molecular mapping; seed germination; plant growth;
 KM plant quality; plant yield; plant breeding; tissue printing; ss.
 XX
 OS *Gossypium hirsutum*.
 XX
 PN US2004123340-A1.
 XX
 PD 24-JUN-2004.
 XX
 PF 12-DEC-2001; 2001US-00021323.
 XX
 PR 14-DEC-2000; 2000US-0255619P.
 XX
 PA (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K L.
 PA (ZIEG/) ZIEGLER T E.
 XX
 PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 PT
 DR WPI; 2004-479808/45.
 XX
 PT New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 XX
 XX
 XX Claim 1; SEQ ID NO 3636; 34p; English.
 XX
 CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN48855/c). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312, Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucleon33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be

CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?docid=US20040123340

XX Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 9.2%; Score 222.4; DB 13; Length 552;

Best Local Similarity 72.2%; Pred. No. 3.3e-54;

Matches 289; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy 2006 AACGACTCCAGAAATGATGCTAGGACTGACAGAAATATGATCCAGTGGCAGAGA 2065
 Db 550 AAGAAATTAAGCTTAAAGATGAGATGCAAGATTTCAAGAGCATTTGTCGAAATGCGAGATA 491
 Qy 2066 TTAAGTCTGTGCTTTTGGGCTGATCAAGCATAGAAATGTTTACAGAGTTTGGATG 2125
 Db 490 TTAATCTGAGGGGTTTGGACCTGATCAAGCATGACCGCTTGAATTAATGCGAGAGTTCTGGATG 431
 Qy 2126 GGGGAATGCTGAAGATTTTGGACCTGACAGAGAGCTGAAATCTGGCAGCTTGGTGT 2185
 Db 430 GTCAAAATGTTGAAGACATGACAGATGTCGACCCGCAAAATGCTCAGCTTGGTGTAT 371
 Qy 2186 ATGATTTATACAGCTTGAAGAACTATCTGTTGACAGTGTGACAGCTCTGACAGATGAAACC 2245
 Db 370 ATGAATTAATGCTTACTGAAATGATGCAATGCAATGCAATGCTTCACTGATGAGCCAGC 311
 Qy 2246 GTGCTCTGTGAGAGCAACTCTGAGAGAGTCTGCTGTATCTGATTTGGTTTATCCAG 2305
 Db 310 GAGCTGATGCAAGCTACTCTGGAAGAAATCCAGCTGTTGATGATGTTTATCATCCAG 251
 Qy 2306 AAAACAATGCTACTGATGTCAGAACTTACACAAAGAAATGCAAGTTTCTGCTCAAGT 2365
 Db 250 AGAACAATGCTCTTAATGTTAACTCTTACACCAAGATATGAGATGTTTGTCCAACT 191
 Qy 2366 CAGGCTGAGAAATACATGAAAGCTCTGTTTTCATCATTA 2405
 Db 190 CAGGCTGAGAAATACATGAAAGCTCTGTTTTCATCATTA 151

RESULT 14

ADJ38255
 ID ADJ38255 standard; cDNA; 545 BP.
 ADJ38255;
 AC 06-MAY-2004 (first entry)
 XX
 DE Placid division-related Arce orthologue cDNA 42.
 XX
 KM prokaryotic type; placid division; Fm2; AR6; AR5; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; gene; ss.
 XX
 OS *Gossypium arboreum*.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.

Db 362 CAATCAACCTTGACACTGCTCTGCTTCACAATTTGATGAAACTGTGAGGAGATCACAC 421
Qy 803 CGGCTATGCTCTTGAGGCTACTTGCGCTTACCGCTTGTGA 842
Db 422 CTCGTGTGTGTTTGAGGCTCTTGCCCTTCTTCTTGATGA 461

Search completed: February 21, 2006, 15:41:27
Job time : 1431 secs

THIS PAGE BLANK (USPS)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 14:22:16 ; Search time 693 Seconds
(without alignments)
7370.114 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagcctgagtcagct.....gctctgtcttgatcataa 2406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.New:*
1: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	46	1.9	556	6 US-09-925-065A-355459	Sequence 355459,
C 2	44.8	1.9	191684	12 US-11-121-086-2	Sequence 2, Appl1
C 3	43.8	1.8	365	12 US-11-043-752-3306	Sequence 3906, Ap
C 4	43.8	1.8	365	12 US-11-043-752-3309	Sequence 3909, Ap
C 5	43	1.8	1281	8 US-10-131-826A-509	Sequence 509, App
C 6	42.8	1.8	479	6 US-09-925-065A-29205	Sequence 29205, A
C 7	42.4	1.8	169495	12 US-11-121-086-61	Sequence 61, Appl
C 8	42.2	1.8	1080000	8 US-10-928-446A-1	Sequence 1, Appl1
C 9	42.2	1.8	1080000	8 US-10-928-446A-181	Sequence 181, App
C 10	42.2	1.8	1080000	8 US-10-928-446A-183	Sequence 183, App
C 11	42.2	1.8	1080000	8 US-10-928-446A-185	Sequence 185, App
C 12	42.2	1.8	1080000	8 US-10-928-446A-187	Sequence 187, App
C 13	42.2	1.8	1080000	8 US-10-928-446A-189	Sequence 189, App
C 14	42.2	1.8	1080000	8 US-10-928-446A-191	Sequence 191, App
C 15	42.2	1.8	1080000	8 US-10-928-446A-193	Sequence 193, App
C 16	42.2	1.8	1080000	8 US-10-928-446A-195	Sequence 195, App
C 17	42.2	1.8	1080000	8 US-10-928-446A-197	Sequence 197, App
C 18	42.2	1.8	1080000	8 US-10-928-446A-199	Sequence 199, App
C 19	42.2	1.8	1080000	8 US-10-928-446A-201	Sequence 201, App
C 20	41.2	1.7	2720	6 US-09-925-065A-74187	Sequence 74187, A

C 21	41	1.7	217	12 US-11-043-752-3461	Sequence 3461, Ap
C 22	41	1.7	262	12 US-11-043-752-3458	Sequence 3458, Ap
C 23	41	1.7	2001	12 US-11-043-752-3454	Sequence 3454, Ap
C 24	41	1.7	2001	12 US-11-043-752-3455	Sequence 3455, Ap
C 25	40.4	1.7	687	6 US-09-925-065A-69053	Sequence 69053, A
C 26	40.4	1.7	7764	12 US-11-136-527-3891	Sequence 3891, Ap
C 27	40.2	1.7	1363	6 US-09-925-065A-74546	Sequence 74546, A
C 28	40.2	1.7	31028	8 US-10-829-826B-21	Sequence 21, Appl
C 29	40.2	1.7	31028	8 US-10-829-826B-22	Sequence 22, Appl
C 30	40.2	1.7	31028	8 US-10-829-826B-24	Sequence 24, Appl
C 31	40.2	1.7	31100	8 US-10-829-826B-26	Sequence 26, Appl
C 32	39.8	1.7	244	12 US-11-043-752-3912	Sequence 3912, Ap
C 33	39.8	1.7	1450	12 US-11-140-417-3	Sequence 3, Appl1
C 34	39.8	1.7	63693	8 US-10-995-561-13269	Sequence 13269, A
C 35	39.8	1.7	168516	12 US-11-121-086-3	Sequence 3, Appl1
C 36	39.4	1.6	199321	12 US-11-121-086-10	Sequence 10, Appl
C 37	39.2	1.6	1317	12 US-11-140-417-1	Sequence 1, Appl1
C 38	39.2	1.6	11462	12 US-11-140-417-22	Sequence 22, Appl
C 39	39.2	1.6	120096	12 US-11-121-086-24	Sequence 24, Appl
C 40	39	1.6	523	6 US-09-925-065A-602447	Sequence 602447, A
C 41	39	1.6	10129	12 US-11-044-111-21	Sequence 21, Appl
C 42	39	1.6	153376	12 US-11-121-086-5	Sequence 5, Appl1
C 43	38.8	1.6	1772	12 US-11-167-856-23	Sequence 23, Appl
C 44	38.8	1.6	7402	8 US-10-750-185-27313	Sequence 27313, A
C 45	38.8	1.6	7402	8 US-10-750-623-27313	Sequence 27313, A

ALIGNMENTS

RESULT 1
US-09-925-065A-355459/C
Sequence 355459, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 355459
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(556)
OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-355459
Query Match
Best local Similarity 1.9%; Score 46; DB 6; Length 556;
Matches 76; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
C 136 CTTCTCTCGATCAATTACCTCCGATTCCTCTCTCTCTCTCCACCGCCACC 195
DB 273 CTTC 214
C 196 ACCACGCGACCTGCTCTCTCTGACCATCATATGATGTCGCAACGCAACGTC 255


```

Query Match 1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY      76  CTCGAGCTAGGCACAACAACCTCTTACACACTATCTGTGCTCCGCGAGCAATAGGCGCAACCGT 135
Db      886419  CTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 886478

QY      136  CTCTCTCTCGACTTCAATTTACACCTCCGATTCCTCTCCCTCCCTCTTTCGACACGGCAACC 195
Db      886479  CTCCTCTCTCGACTCTCTCTCTATGCCCCCTCTCTCTTCAACCTCTCTCTCTCTCTCTCTCTCTCTC 886538

QY      196  ACCACCGCACTCTGCTCTCTCTCTGACCAACATATTGATGTCGCA 242
Db      886539  TCTTTTACCTCTCTCTCTCTCTTCTTTCGACACACCCCAACCCCGCCCCCA 886585

RESULT 9
US-10-928-446A-181
; Sequence 181, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION

```

```

1  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
2  TITLE OF INVENTION: VITAL BUDDING
3  FILE REFERENCE: 0274-5785.1US
4  CURRENT APPLICATION NUMBER: US/10/928,446A
5  CURRENT FILING DATE: 2004-08-26
6  PRIOR APPLICATION NUMBER: 60/359,741
7  PRIOR FILING DATE: 2002-02-26
8  NUMBER OF SEQ ID NOS: 202
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 181
11 LENGTH: 1080000
12 TYPE: DNA
13 ORGANISM: Homo sapiens
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: (826985)..(827008)
17 FEATURE:
18 OTHER INFORMATION: full exon 1 range is 826667-827008
19 NAME/KEY: allele
20 LOCATION: (827008)..(827008)
21 OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
22 FEATURE:
23 NAME/KEY: CDS
24 LOCATION: (843242)..(843315)
25 OTHER INFORMATION: exon
26 FEATURE:
27 NAME/KEY: CDS
28 LOCATION: (922549)..(922630)
29 OTHER INFORMATION: exon
30 FEATURE:
31 NAME/KEY: CDS
32 LOCATION: (926021)..(926059)
33 OTHER INFORMATION: exon
34 FEATURE:
35 NAME/KEY: CDS
36 LOCATION: (929123)..(929176)
37 OTHER INFORMATION: exon
38 FEATURE:
39 NAME/KEY: CDS
40 LOCATION: (993104)..(993154)
41 OTHER INFORMATION: exon
42 FEATURE:
43 NAME/KEY: CDS
44 LOCATION: (999547)..(999608)
45 OTHER INFORMATION: exon
46 FEATURE:
47 NAME/KEY: CDS
48 LOCATION: (1000354)..(1000456)
49 OTHER INFORMATION: exon
50 FEATURE:
51 NAME/KEY: CDS
52 LOCATION: (1002118)..(1002284)
53 OTHER INFORMATION: exon
54 FEATURE:
55 NAME/KEY: CDS
56 LOCATION: (1006117)..(1006249)
57 OTHER INFORMATION: exon
58 FEATURE:
59 NAME/KEY: CDS
60 LOCATION: (1007860)..(1008036)
61 OTHER INFORMATION: exon
62 FEATURE:
63 NAME/KEY: CDS
64 LOCATION: (1010940)..(1011014)
65 OTHER INFORMATION: exon
66 FEATURE:
67 NAME/KEY: CDS
68 LOCATION: (1018160)..(1018291)
69 OTHER INFORMATION: exon
70 FEATURE:
71 NAME/KEY: CDS
72 LOCATION: (1018800)..(1018919)

```

Query Match	1.8%	Score 42.2;	DB 8;	Length 1060000;
Best Local Similarity	53.3%;	Pred. No. 6.5;		
Matches 89;	Conservative 0;	Mismatches 78;	Indels 0;	Gaps 0
76	CTCCGACGTAAGCCACAACCTTACACACTATCTGCTCCGCCAGCAAAATGGCCGACCT	135		

```

Db      886419  CTCCCGCTCCCTCCGCCCCCCCCCTCCCTCTCTCCCTCTCCGCCCCCTCTC 886478
Oy      136  CTCTCTCCGACTTCAATTACCTCCGATTCCTCTCTCTCTCTGCGCACCGGCACC 195
Db      886479  CTCTCTCCGACCTCCGATCTGCCCCCTCTCTTCCACTCTCTCTCTCCACCTCCCC 886538
Oy      196  ACCACCGGACCTCTGCTCTCTGCGCACCATCTATTGATTCGCCCCGA 242
Db      886539  TCTTTTACCTCTCTCTCTCTTTCTTGACACACCCACCCCGCCCCCA 886585

RESULT 10
US-10-928-446A-183
; Sequence 183, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
; TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
; TITLE OF INVENTION: VIRAL BUDDING
; FILE REFERENCE: 0274-5785.1US
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/359,741
; PRIOR FILING DATE: 2002-02-26
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 183
; LENGTH: 1080000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (826985)..(827017)
; FEATURE:
; OTHER INFORMATION: full exon 1 range is 826667-827008
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (843242)..(843315)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (922549)..(922630)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (926021)..(926059)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (929123)..(929176)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (993104)..(993154)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (999547)..(999608)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1000354)..(1000456)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1002118)..(1002284)
; FEATURE:
; OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1006117)..(1006249)
; FEATURE:
; OTHER INFORMATION: exon
; FEATURE:

```

[illegible]

[illegible]

```

OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018160)..(1018291)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018800)..(1018919)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1020028)..(1020225)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1026659)..(1026736)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1028113)..(1028167)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1034316)..(1034374)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1041390)..(1041455)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1043121)..(1043350)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1044868)..(1044989)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1047519)..(1047589)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1050296)..(1050391)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1060368)..(1060441)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1062648)..(1062708)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1064561)..(1064620)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1066207)..(1066314)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1067768)..(1067864)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1068609)..(1068681)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1073289)..(1073388)
FEATURE:

```

```

OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-189
Query Match 1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 76 CTCGACGTAGCACACACCTCTACACTATCTGCTCCGACGACAAATGGACCGT 135
DB 886419 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886478
QY 136 CTCTCTCTCGACTTCAATTTACACTCTCGATCTCTCTCTCTCTCTCTCTCT 195
DB 886479 CTCCTCTCTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886538
QY 196 ACCACGCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
DB 886539 TCCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886585

RESULT 14
US-10-928-446A-191
Sequence 191, Application US/10928446A
Publication No. US20050277123A1
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
TITLE OF INVENTION: VIRAL BUDDING
FILE REFERENCE: 0274-5785.1US
CURRENT APPLICATION NUMBER: US/10/928,446A
CURRENT FILING DATE: 2004-08-26
PRIOR APPLICATION NUMBER: 60/359,741
PRIOR FILING DATE: 2002-02-26
NUMBER OF SEQ ID NOS: 202
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 191
LENGTH: 1080000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (999562)..(999608)
FEATURE:
OTHER INFORMATION: full length exon 7 range is 999547-999608
FEATURE:
NAME/KEY: CDS
LOCATION: (1000354)..(1000456)
FEATURE:
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1002118)..(1002284)
FEATURE:
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1006117)..(1006249)
FEATURE:
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1007860)..(1008036)
FEATURE:
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1010940)..(1011014)
FEATURE:
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018160)..(1018291)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018800)..(1018919)
OTHER INFORMATION: exon
FEATURE:

```



```

1 NAME/KEY: CDS
2 LOCATION: (1020028) .. (1020225)
3 OTHER INFORMATION: exon
4 FEATURE:
5 NAME/KEY: CDS
6 LOCATION: (1026659) .. (1026736)
7 OTHER INFORMATION: exon
8 FEATURE:
9 NAME/KEY: CDS
10 LOCATION: (1028113) .. (1028167)
11 OTHER INFORMATION: exon
12 FEATURE:
13 NAME/KEY: CDS
14 LOCATION: (1034316) .. (1034374)
15 OTHER INFORMATION: exon
16 FEATURE:
17 NAME/KEY: CDS
18 LOCATION: (1041390) .. (1041455)
19 OTHER INFORMATION: exon
20 FEATURE:
21 NAME/KEY: CDS
22 LOCATION: (1043121) .. (1043350)
23 OTHER INFORMATION: exon
24 FEATURE:
25 NAME/KEY: CDS
26 LOCATION: (1044868) .. (1044989)
27 OTHER INFORMATION: exon
28 FEATURE:
29 NAME/KEY: CDS
30 LOCATION: (1047519) .. (1047589)
31 OTHER INFORMATION: exon
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: (1050296) .. (1050391)
35 OTHER INFORMATION: exon
36 FEATURE:
37 NAME/KEY: CDS
38 LOCATION: (1060368) .. (1060441)
39 OTHER INFORMATION: exon
40 FEATURE:
41 NAME/KEY: CDS
42 LOCATION: (1062648) .. (1062708)
43 OTHER INFORMATION: exon
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: (1064561) .. (1064620)
47 OTHER INFORMATION: exon
48 FEATURE:
49 NAME/KEY: CDS
50 LOCATION: (1066207) .. (1066314)
51 OTHER INFORMATION: exon
52 FEATURE:
53 NAME/KEY: CDS
54 LOCATION: (1067768) .. (1067864)
55 OTHER INFORMATION: exon
56 FEATURE:
57 NAME/KEY: CDS
58 LOCATION: (1068609) .. (1068681)
59 OTHER INFORMATION: exon
60 FEATURE:
61 NAME/KEY: CDS
62 LOCATION: (1073289) .. (1073388)
63 FEATURE:
64 OTHER INFORMATION: full length exon 30 range is 1073289-1075279
65 JS-10-928-446A-191

```

	Query Match Similarity	1.8%	Score 42.2;	DB 8;	Length 1080000;
	Best Local Similarity	53.3%	Pred. No. 6.5;		
	Matches	89;	Conservative	0;	Mismatches 78; Indels 0; Gaps 0;
Oy	76	CTCGAGGTGACCAACTTCTATCTATCTGCCGACGAATGGCCGACCGT	135		
db	886419	CTCCCCCTCCTCTCCGCCCTCCCTCCTCTCCTCTCCTCTCCTCCGCCCCCTTC	886478		

QY 136 CTTCCTCCGACTTCAATTACACCCGATTCCTCTCTCTCTCTGCGACACCGGACG 195
Db 886479 CTCCTCCACACCTCTACTGCGCCCTCTCTTTCACCTCTCTCTCGACACCTCCCC 886538
QY 196 ACCACCGGCACTCTGTCTCTCTGCGACGATATATGATCTGCCGA 242
Db 886539 TCTCTTACTCTCTCTCTCTCTTCGACACCCGACCCCGCCCCA 886585

```

RESULT 15
US-10-928-446A-193
; Sequence 193, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
; TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
; TITLE OF INVENTION: VIRAL BUDDING
; FILE REFERENCE: 0274-5785.1US
; CURRENT APPLICATION NUMBER: US/10/928,446A
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/359,741
; PRIOR FILING DATE: 2002-02-26
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 193
; LENGTH: 1080000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (999562)..(999608)
; FEATURE:
; OTHER INFORMATION: full length exon 7 range is 999547-999608
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1000354)..(1000456)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1002118)..(1002284)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1006117)..(1006249)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1007860)..(1008036)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1010940)..(1011014)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1018160)..(1018291)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1018800)..(1018919)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1020028)..(1020225)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1026659)..(1026736)
; OTHER INFORMATION: exon
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1028113)..(1028167)

```

Search completed: February 21, 2006, 14:58:05
Job time : 699 secs

```
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1034316)..(1034374)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1041390)..(1041455)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1043121)..(1043350)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1044868)..(1044989)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1047519)..(1047589)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1060368)..(1060441)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1062648)..(1062708)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1064561)..(1064620)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1066207)..(1066314)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1067768)..(1067864)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1068609)..(1068681)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1073289)..(1073388)
FEATURE:
OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-193
```

Query Match 1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
QY 76 CTCGACGTAGCAGACACACTCTCACTATCTGCTCCGACAGCAATGAGCGACCGT 135
DB 886419 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886478

QY 136 CTTCTCTCGAGTTCAATTTCACTCGAATCTCTCTCTCTCTCTCTCTCTCTCTCT 195
DB 886479 CTCCTCTCTCACTCTCACTGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 886538

QY 196 ACCACCGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
DB 886539 TCCTTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886585
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 21, 2006, 12:51:25 : Search time 10553 Seconds
(without alignments)
4314.565 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGIGLSPFCRLP.....YEVFWKSGWKITGSVTLAS 801

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=rlh
-Q=/abs/ABSSWEB.spool/US10600070/runatc_21022006_112358_8614/app_query.fasta-1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blcsm62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss03h
-USER=US10600070 @CCN 1.1_4939 @runatc_21022006_112358_8614 -NCPU=6 -ICU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV -TIMEOUT=120
-MARN -TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_da.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_srs.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4052	99.7	2438	15	AY221469 Arabidops
2	4051	99.7	2436	15	AY150490 Arabidops
3	4051	99.7	2637	15	AY091075 Arabidops

4	3837.5	94.4	3668	15	AY221468	AY221468 Arabidops
5	3825.5	94.2	85791	15	AB016888	AB016888 Arabidops
6	3823.5	94.1	3664	15	AY221467	AY221467 Arabidops
7	1540.5	37.9	110000	15	AP008208_011	Continuation (12.0)
8	1540.5	37.9	150452	15	AP004885	AP004885 Oryza sat
9	1241	30.5	133779	14	AC158210	AC158210 Medicago
10	1222	30.1	184752	14	AC155597	AC155597 Zea mays
11	702	17.3	110000	14	CR954204_0	CR954204 Oetereococ
12	497	12.2	110000	1	BA000019_32	Continuation (33.0)
13	497	12.2	110000	1	BA000045_36	Continuation (34.0)
14	449	11.1	110000	1	BA000045_36	Continuation (37.0)
15	446	11.0	2469	1	AF421196	AF421196 Synecoco
16	446	11.0	110000	1	AP008231_22	Continuation (23.0)
17	436.5	10.7	110000	1	BA000022_23	Continuation (24.0)
18	431	10.6	2857	15	AY074283	AY074283 Arabidops
19	425	10.5	110000	1	BA000038_07	Continuation (8.0)
20	357.5	8.8	110000	1	CP000095_07	Continuation (8.0)
21	357.5	8.8	110000	1	CP000095_07	Continuation (9.0)
22	351	8.6	550	10	DQ022571	DQ022571 Beta vulg
23	344.5	8.5	348986	1	BX572093	BX572093 Prochloro
24	340	8.4	349213	1	BX565693	BX565693 Synecoco
25	323	7.9	349323	1	BX572096	BX572096 Prochloro
26	318	7.8	300143	1	AE017165	AE017165 Prochloro
27	299	7.4	117818	14	AC160013	AC160013 Medicago
28	229.5	5.6	110000	15	AP008210_345	Continuation (346)
29	229.5	5.6	164943	15	OSJG000052	AL606457 Oryza sat
30	227.5	5.6	104146	14	OSJG000052	Continuation (24.0)
31	220	5.4	110000	1	AP008231_23	AP000419 Arabidops
32	185	4.6	81875	15	AP000419	Continuation (29.0)
33	167.5	4.1	110000	1	AY596297_28	AB050629 Bacillus
34	155.5	3.8	41599	6	E63260	E63260 Gene encodi
35	155.5	3.8	41599	6	E63260	Continuation (6.0)
36	151.5	3.7	41884	1	BAM575642	BAM575642 Bacillus
37	150	3.7	11237	1	AE007907	AE007907 Agrobacte
38	150	3.7	12027	1	AE008959	AE008959 Agrobacte
39	149	3.7	110000	1	AP006878_09	Continuation (10.0)
40	148.5	3.7	6270	6	126720	Sequence 25
41	148.5	3.7	80841	6	CO975935	CO975935 Sequence
42	148.5	3.7	122863	1	BAM575417	BAM575417 Bacillus
43	148.5	3.7	122863	1	BAM576102	BAM576102 Bacillus
44	148	3.6	46513	1	AY137375	AY137375 Bacillus
45	146.5	3.6	110000	1	CR555306_05	Continuation (6.0)

ALIGNMENTS

RESULT 1
AY221469
LOCUS
DEFINITION
Arabidopsis thaliana division protein (ARC6) mRNA, complete cds;
nuclear gene for chloroplast product.
ACCESSION
AY221469
VERSION
AY221469
KEYWORDS
GI:33436353
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 2438)
Vittha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Exp,H. and
Osteryoung,K.W.
Osteryoung,K.W.
TITLE
ARC6 is a J-Domain Plastid Division Protein and an Evolutionary
Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL
PUBMED
12897262
Plant Cell 15 (8), 1918-1933 (2003)
AUTHORS
2 (bases 1 to 2438)
Vittha,S., Koksharova,O., van Exp,H., Froehlich,J.E. and
Osteryoung,K.W.
Direct Submission
TITLE
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State
University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
Location/Qualifiers

```

source
1. .2438
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Wassilewskija"
1. .2438
/gene="ARC6"
/ncbi="synonym: At5g42480"
2. .2407
/gene="ARC6"
/codon_start=-1
/product="division protein"
/protein_id="A018646.1"
/db_xref="GI:33436354"
/translation="MEALSHVIGLSPOLCRUPATTKLRSHNTSTTICASKWAD
RLUSDNFTSDSSSPATRTTATVSLPSIDRPRNRPRIPIIDYQVLGNATHLT
DGIIRAEARVSKPQPFESDDALISRQILQACETLSNPRSRNRBGLDDEGAT
VITDVPMDKPGALCVLOEGETEIVLRVGEALLKRLPSFKQDVLVWALAFDVS
RDMAIDPDPIFGVEFEALKLQEGASLAPGLRAQIDETLEIIPRYVLEILG
LPGLDDYAAKRLNGLSGVNIIIMSVGGGSAVLVGLTRKFPNNAEFLMTAAEOVL
PVATPSNIPAKESFEYVEVALVAQAFIKGKPHLLDADQFOLOAKVMAMEIPAM
LVDRNNMEIDFELERGLCALIGKDECMMLGDSSEDSQYRNPALVERVLNSRD
DNDLPEGLCTLETLGAVFPRPRDKKFKLGDYRDDPWLSTLEREVVQGSPL
AAATMARIGAEHVKAAMQALQKVFPSRYTDNNSAEPKQVETFSVDPNGVRD
GERGVFLAEVRPSENPEITDVAIRAGVSSSDVETTVMSVAMLKESVKKLAAGV
AIGLISFSOKYFLKSSSFORKDMVSMESDVATIGSVADSEALPRMDATANI
VSKWKIKSLAPGPDHRIEMLPRVLDGRMLKIWDRAAETRAQGLVADYTLKLSVS
VTVSADGTRALVATIEESACLSDLVHPENNAIDVRYTTRIVFVSKSMKITBSGV
LMS"
misc_difference
474
/gene="ARC6"
/ncbi="results in Gly to Glu substitution"
/replace="a"
misc_difference
1831
/gene="ARC6"
/ncbi="results in Glu to Asp substitution"
/replace="t"

ORIGIN

Alignment Scores:
Pred. No.: 1.31e-248 Length: 2438
Score: 4052.00 Matches: 799
Percent Similarity: 99.9% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.7% Indels: 0
Gaps: 0
DB: 15

US-10-600-070b-2 (1-801) x AY221469 (1-2438)
QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyAspGluPro 20
Db 2 ATGAAAGCTCTGAGTCACGTGGCATTTGCTCTCCCATTCCTCAATTATGCCATTACCA 61
QY 21 ProAlaThrThrLyLeuAArgSerHisAsnThrSerThrIleCySerAlaSer 40
Db 62 CCGGCGACGACAAAGCTCCGACGTAGCCACACACCTCTAACACTATCTGCTCGGCAGC 121
QY 41 LysThrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
Db 122 AAATGGGCGACCGCTCTCTCTCCGACTTCAATTACCTCCGATTCCTCTCTCTCC 181
QY 61 PheAlaThrAlaThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
Db 182 TTGGCCACCGCCACACACCGCCACTCTGCTCTCTCCACCACTATTTGATGCTGCC 241
QY 81 GluArgHisValProIleProIleAspPheTyrgIleValIleuGlyAlaGlnThrHisPhe 100
Db 242 GAAGCGACCGTCCCACTCCCATTTGATTTCTACCAAGGATTTAGAGAGCTCAAAACACTTC 301
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
Db 302 TTAACCGATGGAATCAGAAAGACATTCGAGCTAGGCTTTCGAAACCGCGCAATTCGGT 361

```

```

QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
Db 362 TTCAGCGACGACGCTTATCAGCCGACGACAGATTCCTCAAGCTGCTTGCAGAACTTCG 421
QY 141 SerAsnProArgSerArgArgGluTyrgArgGlnGlyLeuLeuAspAspGluGluAlaThr 160
Db 422 TCTAATCCTCGCTAGAAAGAGTACATGAAAGGCTCTTGTATGATGAAAGAGCTTACA 481
QY 161 ValIleThrAspValProThrAspLysValProGlyAlaLeuCyAspValIleuGlnGly 180
Db 482 GTCATCATGATTCCTTCCTGGATTAAGTTCCTCGGGGCTCTGCTATATGCAAGAAAGGT 541
QY 181 GlyIleThrGluIleValLeuArgValGlyGlyAlaLeuLeuLysGluArgLeuProLys 200
Db 542 GGTGAACATGAGATGATGCTTCGCGGTGGTGAAGGCTCTCTTAAGAGAGGTTGCTTAAG 601
QY 201 SerPheLysGlnAspValIleuValIleuValIleuAlaLeuAlaPheLeuAspValIleArgAsp 220
Db 602 TCGTTTAAACCAAGATGTGTTTAAATTAGGCGCTTCCGCTTCTCGATGCTCGAGGGAT 661
QY 221 AlaMetAlaLeuAspProPheIleThrGlyTyrgLysPheValGluGluAlaLeu 240
Db 662 GCTATGGCATTTGATCCACTGATTTTATTAAGTTATGAGTTTGTGAGAAAGCTTGG 721
QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
Db 722 AAGCTTTTACAGAGGAGAGACAGCAAGTACCTTGCACCGGATTTACGTGCAAAATTGAT 781
QY 261 GluThrLeuGlnGluIleThrProArgTyrgValLeuGlnLeuLeuGlyLeuProLeuGly 280
Db 782 GAGACTTTGAAAGATCACTCCGCGTATGTTGTTGAGCTACTTGGCTTACCGCTTGGT 841
QY 281 AspAspTyrgAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
Db 842 GATGATTAAGCTCCGAAAGACAAATGTTTAAAGCGGTGGCGAAATTTTGTGCTGT 901
QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
Db 902 GTTGAGAGAGGTGACATCAGCTCTTGTGGGGGTTTGACCCGCTAGAAATTATGAAT 961
QY 321 GluAlaPheLeuAspMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
Db 962 GAGCGGTTTTTACAAATGACAGCTGCTGACAGATGATCTTTGTAGCTTACCCCAAGC 1021
QY 341 AsnIleProAlaGluSerPheGluValTyrgLysValAlaLeuAlaLeuValAlaGlnAla 360
Db 1022 AATATTCACAGACAGCTATTTGAAGTTTACAGAGTTCGATTCCTCTGCTGCTCAAGCT 1081
QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGln 380
Db 1082 TTTATTTGTTAAGACCACTTTTACAGATGCTGATTAAGCAATTCACAGCAACTTCAG 1141
QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrgAspPheArgAsnAsnTrp 400
Db 1142 CAGGCTTAAGGTATGCTATGAGATTCCTGCATGTGTATATACACGGAATTAATTGG 1201
QY 401 GluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
Db 1202 GAGATGACTTCGCTCTAGAAAGGACCTGTGCACCTGTTTAAAGCAAAAGTTGATGAA 1261
QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrgAsnProAlaIleVal 440
Db 1262 TTCCGTTATGTGTGTGGCTTAGACAGTGAAGATTCACAAATATGGAATCCACGCTATTGTG 1321
QY 441 GluPheValIleuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
Db 1322 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGGAATATGCAA 1381
QY 461 LeuLeuGluThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
Db 1382 TTGTTGAAACCTGGTGTGGAGGGGTGTCTTCTCTAGGTTCAGAGACCAAAAGATAAA 1441
QY 481 LysPheLysLeuGlyAspTyrgTyrgAspAspPheMetValLeuSerTyrgLeuGluVal 500

```

D	b		1442	AATTTTAACCTCGGGGCACTACTAGATATCCATTAGTGTTCGATTAAGTGAAAGAGCTG	1501
O	y		501	GUUAlValGInGIySerProLeuAlaaIaaIatHMeCaIaArgIleGIYalAGlu	520
D	b		1502	GAGGAGATTGACGGGGTTCCTCTTAGCTCTGCTGCACCTAATGGAAGAATTGGAGCCGAG	1561
O	y		521	Ht8VAllybAlaSerAlaMetGlNaIaleuGlnlysValPheProSerArGTyrThrAsp	540
D	b		1562	CATGTGAAGGTAGTGTCTATCAGGACCTGCAGAAAGTTTTCTCCCGCTATAACAGT	1621
O	y		541	ArgAnSerAlaGluProLyAspValGInGIuTrValIPheSerValAspProValGIY	560
D	b		1622	AGAACTCGCTGAACCAGAGAGATGCGAACAGACAGTGTTAAGTATGATCCCTGTGGT	1681
O	y		561	AenAnValGIYarGspGIYGIuProGIYalPheIleaGIuaIaValaIargProser	580
D	b		1682	AACAATGATGAGCCGGATGATGTGAGCTGTGTCTTATGTGCAAGAGCTGAAGCCCTCT	1741
O	y		581	GIuBnPhelGIuTrnAenAPTyRAlaIleaArgAlaGIYalSerGIuSerSerValAsp	600
D	b		1742	GAAAACTTGTGAACATAAGATTAATGCCAATTGCAGCTGGGGTCTCAGACAGTACGGTTGAT	1801
O	y		601	GIuHrThrValIGlumeSerValaIaaPhmetLeuysGIuaIaSerValIyleu	620
D	b		1802	GAAACTACTGTGTAATGTCGGTGTCTAGATGTAAAGAGAGCGCAAGGTAAAGTCTTA	1861
O	y		621	AlaAlaGIYalAlaIleGIYleuIeSerLeuPheserGlnIyTYrPheLeuLysSer	640
D	b		1862	GCTGCTGGTGGCAATTGCATGATTCACTGTTCACCGCAAGAGTATTTCTTTAAAAGC	1921
O	y		641	SerSerSerPheGlnArgLYaSPMeValSerSerMetGIuSerAspValaIaThrIle	660
D	b		1922	AGCTCATCTTTTCCAACGACAGATTAAGTTCCTCTAATGAAATCGATGTGTGCTACATTA	1981
O	y		661	GIYserValaArgAlaaspAspSerGIuNlaIeuProArgMetARPaIaArgThrAlaGIu	680
D	b		1982	GGGTCAGTCAGAGCTGACGATTCAGAACCACTTCCCAGAAATGATGATGCTGAGACTGCAGAG	2041
O	y		681	AenIleValSerLYSTRPGLnllysIleySerLeuAlaPheGIYProasPHIaArgIle	700
D	b		2042	AATATAGTATCCAGAGTGGCAGAAAGATTAACTCTGCGCTTTTGGGCCGTGACACCGCATTA	2101
O	y		701	GIuMetLeuProGIuValIleuaspGIYArgMetLeuysIIetTPThrAspArgAlaIa	720
D	b		2102	GAAATGTATTACCAAGAGTTTGGATGGCGGAATGCTGAAGAGATTGGACTGACAGAGCAGCT	2161
O	y		721	GIuHrAlaGlnLeuGIYleuValTYrAspTYrThrLeuLeuLYbleuSerValAspSer	740
D	b		2162	GAAACTGGCAGAGCTGGGTGGTTATATATATACACTGTGAACATATCTGTTGACAGT	2221
O	y		741	ValIThrValISerAlaaspGIYTHrArgAlaIeuValGIuaIaTHrLeuGIuSeraIa	760
D	b		2222	GTCACAGTCTCACAGATGAGAAACCCGTCTCTGTGGAGAACCACTCTGAGAGATGTGCT	2281
O	y		761	CysLeuSerAspLeuValIaIsProGIuAnbanAnaIthrAspValaIargTYrThrThr	780
D	b		2282	TGTGTATCTGAATTTGGTTCATCCAGAAACAATGTACTGATGTACAGAACTTACACAACA	2341
O	y		781	ArgTYrGIuValPheTPSerLYsSerGIYTRPLysIIethrGIuGIYSerValIleuAla	800
D	b		2342	AGATACGAAGATTTCTGTGCTCCAAAGTCAGAGGTGGAANAATCACTGAAGGCTCTGTGTTCTTGA	2401
O	y		801	Ser 801	
D	b		2402	TCA 2404	

RESULT 2
LOCUS AY150490 2436 bp mRNA linear PLN 23-SEP-2002

DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.

ACCESSION AY150490

VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AY150490.1	GI:23297711	FLI CDNA	Arabidopsis thaliana (thale cress)	Arabidopsis thaliana Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	
1 (bases 1 to 2436)					
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.					
Unpublished					
2 (bases 1 to 2436)					
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.					
Direct Submission					
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA					
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.					
The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinm,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.					
Yamada,K. (SSP/PBGC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PBGC) contributed equally to this work as PIs.					
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.					
Location/Qualifiers					
1..2436					
/organism="Arabidopsis thaliana"					
/mol_type="mRNA"					
/db_xref="taxon:3702"					
/chromosome="5"					
/clone="U19395"					
/ecotype="Columbia"					
/note="This clone is in pUNI 51."					
1..2436					
/gene="At5g42480"					
1..2406					
/gene="At5g42480"					
/codon_start=1					
/evidence=experimental					
/product="unknown protein"					
/protein_id="AA12907.1"					
/db_xref="GI:23297712"					
/translation="MALSVAGIGLSPFCRLPPATTKRKRSHTSTTCSASKAND RLSDVFTSDSSSSPAITATTATVTPPSIDRPRHPVIPIDFYQVGAQHFTIT DGRRAEARVSRKPQPGFDALISRQILQACETLSNPRSREYNBLDDEKAT VTRDPKAVGALCVLOEGGEITEIVLRVEALKERLPKSGKODVVLVWALAFLDVS RDMAADPDPIGTGYEFARLKLIKEGASSLAPDLRAOIDETLEISITPVYLELG LPLADSDIAAKRLINGSVRNIINSVGKGASALVCGITREKPNEMAFIRMTABOUNL FVATPSVIAPSEFEVVALVAQAATGKKPHLLDADKQFOQLQDAKMAHEIPPM					

LVDTRNNMEIDFGLERGLCALLIGKVDCECMWLGJLSEDSQYRNPALVEFVLNSNRD
DNDLPLGLCLLETWLAGVVEPRFRDKKFLGIDYDDPMLSTLERFEVQSGPL
AAAAAARIGAEHVKAAMQKQFPSSAPKQVQETVSVDVGNVQSRD
GERGVFLAEVRPSENDETVAIRAGVSSVDETVEASVDMKEASVKILIAAGV
AIGLILFPOKXVPLKSSSPORDMVSMSDVATIGSVADNSELPRMDATANI
VSRKQIKSLAFEPDRHTEMLPEVLDGRMKITTDRAETQGLVYDITLKLSDS
VTVSADGTRALVEATLEESACLSDLVHPENNADIVRTYTTTVEFWKSGMKITGESV
LAS"

3'UTR
2407..2436
/gene="Ac5g42480"
misc_difference 2433.2434
/gene="Ac5g42480"
/note="compared to genomic sequence"
/replace="ctg"

ORIGIN

Alignment Scores:

Pred. No.:	1,52e-248	Length:	2436
Score:	4051.00	Matches:	799
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	2
Query Match:	99.7%	Indels:	0
DB:	15	Gaps:	0

US-10-600-070b-2 (1-801) x AY150490 (1-2436)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
DB 1 ATGGAGCTTGATGATCGTCGCGCATTCCTCTCCCATTCCTCAATTTATGCCATTTACCA 60
QY 21 ProAlaThrThrIleuArgSerHisAsnThrSerThrIleCySerAlaSer 40
DB 61 CCGGCGACGACAAAGCTCCGACGACGACCAACCTCTCAACTATCTCGCGCAGC 120
QY 41 LysETPAlaAspArgLeuLeuSerHisAspPheAsnThrSerSerSerSerSer 60
DB 121 AAATGGCCGACCGCTCTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
DB 181 TTGGCCACGCGCACCGACCGCACCTCTCTCTCTCCGACCATTTATTTATTCGTC 240
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100
DB 241 GAACGCGACGTCCTCCATCCCATTTGATTTTCAACAGGATTTAGAGGCTCAACACATTC 300
QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProGlnPheGly 120
DB 301 TTAAACGATGGAATCAGAAAGACATTCGAAGCTAGGGGTTTCGAAACCGCGCAATTCG 360
QY 121 PheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCySGluThrLeu 140
DB 361 TTACGACACACGCTTTAATCAGCCGAGACAGATTTTCAAGCTGCTTCCGAAACCTCTG 420
QY 141 SerAspProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspAspGlnGlnAlaThr 160
DB 421 TCTAATCTCTGCTGTAAGAAGATACATGAAGGCTCTTCTGATGATGAAGAAGCTACA 480
QY 161 ValIleThrAspValProTTPAspLysValProGlyAlaLeuCyValLeuGlnGlnGly 180
DB 481 GTCATCACTGATGTTCTCTGGGATAGGTTCTGCTGCTCTGCTGATATGCAAGAGGT 540
QY 181 GlnGluThrGlnIleValLeuArgValGlyGlnAlaLeuLeuLysGlnArgLeuProLys 200
DB 541 GGTGAGCTGAGATGATGTTCTGCGGTTGGAGGCTCTGCTTAAAGGAGAGGCTTAAAG 600
QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
DB 601 TCGTTTAAAGCAAGATGATGTTTAAATGATGAGCGCTTCGTTTCTCGATGCTCGAGGAT 660
QY 221 AlaMetAlaLeuAspProPheAspPheIleThrGlyTyrGlnPheValGlnGlnAlaLeu 240
DB 661 GCTATGCGCATTTGATCCACTGATTTTATTAATGCTGTTATGATGTTTGTGGAAGCTTTG 720

QY 241 LysLeuLeuGlnGlnGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
DB 721 AAGCTTTTACAGGAGGAGGAGGAGGAGGATGACCTTGGCACCCGATTTAGTCACAAATTTGAT 780
QY 261 GluThrLeuGlnGlnIleThrProArgTyrValIleLeuGlnLeuGlnIleLeuProLeuGly 280
DB 781 GAGACTTTGGAAAGATGATCACTCCGCTTATGCTTGGACCTTACTTGGCTTACCGCTTGGT 840
QY 281 AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
DB 841 GATGATTTACGCTCCGAAAGACCTAATGCTTTAAGCGGTGCGGAAATTTTGTGCTCT 900
QY 301 ValGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGlnLysPheMetAsn 320
DB 901 GTTGAGAGAGGTGAGACATCAGCTCTGTTGGGGGTTTACCCGCTGAGAAGTTATGAAT 960
QY 321 GluAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
DB 961 GAGGCTTTTTCGAAATGACAGCTGTCGACAGGTTGATCTTTTGTAGCTACCCCAAC 1020
QY 341 AsnIleProAlaGluSerPheGluValTyrGlnValAlaLeuAlaLeuValAlaGlnAla 360
DB 1021 AATATTCAGCAGATCAATTTGAACTTTACGAAGTTGCACTTGCTTTGCTGCTCAAGCT 1080
QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
DB 1081 TTTATTTGTAAGAAAGCACACCTTTTACAGATGCTGATACAAATTCAGCAACTTACAG 1140
QY 381 GlnAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAsnTrp 400
DB 1141 CAGGCTAAGGTATAGCTATGAGATCTCGCATTTTATGATACAGGAAATTAATTTG 1200
QY 401 GluIleAspPheGlyLeuGlnArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGln 420
DB 1201 GAGATGACTTCGCTTTAAGAAAGGAGACTCTGTGCACTGCTTTAAGGCAAGTTGATGA 1260
QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
DB 1261 TGGCGTATGAGTGGGCTTTAGCACTGAGGATTCACAAATTAAGAAATTCAGCTAATG 1320
QY 441 GluPheValLeuGlnAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
DB 1321 GAGTTTGTGTTGGAAATTCAAATCGTATGACAAAGATGATCTCTCGACTGATGCAAA 1380
QY 461 LeuLeuGluThrTrpLeuAlaGlyValIlePheProArgPheArgAspThrLysAspLys 480
DB 1381 TTGTTGAAACCTGGTTGGCAGGGGTTGCTTTCTTCTAGGTTCAAGACACCAAGATTA 1440
QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGlnArgVal 500
DB 1441 AATATTTAACTGGGGGACTTATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 501 GluValValGlnGlySerProLeuAlaAlaAlaIleThrMetAlaArgIleGlyAlaGln 520
DB 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValIlePheProSerArgTyrThrAsp 540
DB 1561 CATGTAAAGCTAGTCTATGACGAGCACTGCAAGAAATTTTCTTCCGCTTATACAT 1620
QY 541 ArgAsnSerAlaGluProLysAspValGlnGlnThrValPheSerValAspProValGly 560
DB 1621 AGAAATTCGGCTGAACCAAGATGTCGAAGAGACAGTGTATAGTATGATCCTGTTGGT 1680
QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGlnAlaValAspProSer 580
DB 1681 AACCAATGATGAGCGCTGATATGAGCTGCTGCTTATATGCGAAGACCTGTAAAGACCTCT 1740
QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
DB 1741 GAATACTTGAACCTAATGATTTATGCAATTCGAGCTGGGGTCTTCAAGAGATGACCTTGA 1800

Qy	601	GlutThrValGluMetSerValAlaAspMetLeuIleGluAlaSerValIleLeu	620
Db	1801	GAAGACATCTTGAAGATGTCGTTGCTGATATGTTAAAGAGCAATGTAAGATCTTA	1860
Qy	621	AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleYrPheLeuIleSer	640
Db	1861	GCTGCTGCTGTCGCAATGCACTGATTCATCTGTCACCCAGAAATCTTTCTTAAAGC	1920
Qy	641	SerSerSerPheGlnIleGlyValAspMetValSerSerMetGluSerAspValAlaThrIle	660
Db	1921	AGCTCATCTTTTCAACCAAGAAATGATGTTCTTCTTGAATCTGATGTCCTACCATTA	1980
Qy	661	GlySerValAlaAlaAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu	680
Db	1981	GGGTCACTGACAGCTGACATTCAGAACCACTTCCAGATGATGCTAGAGCTGACAG	2040
Qy	681	AsnIleValSerIleYrPheGlnIleIleYleSerLeuAlaPheGlyProAspHisArgIle	700
Db	2041	AATATAGTATCCAAAGTGCAGAAAGATTAAGTCTCTGCTTTTGGGCTGATCAGCCGATA	2100
Qy	701	GluMetLeuProGluValLeuAspGlyArgMetLeuIleIleTyrPheAspAlaAla	720
Db	2101	GAAGATGTACCAAGGTTTGGATGCGCGAATGCTGAAGATTTGGACTGACAGAGCAGCT	2160
Qy	721	GluThrAlaGlnLeuGlyLeuValIleYrAspTyrThrLeuLeuIleSerValAspSer	740
Db	2161	GAAGATCGGACGCTGGGTTGCTTATGATTAATACCTGTTGAACATATCTGTTGACAGT	2220
Qy	741	ValThrValSerAlaAspGlyThrArgAlaLeuValAlaThrLeuGlnIleGluSerAla	760
Db	2221	GTGACAGTCTCAGACAGATGAAACCGCTGCTGCTGAGAACCACTGAGAGAGTCTGCT	2280
Qy	761	CysLeuSerAspLeuValHisPheGluIleAsnAlaThrAspValAlaGlyThrIleThr	780
Db	2281	TGCTATCTGATTTGGTTCTTCCAGAAACATGCTATGTCAGAACCTTACACACA	2340
Qy	781	ArgTyrGluValPheTyrSerIleYrTyrPheIleThrGluGlySerValLeuAla	800
Db	2341	AGATACGAAGTTTCTGCTGCAAGTCCAGGTTGAAATATCACTGAAGGCTCTGTTCTTGA	2400
Qy	801	Ser 801	
Db	2401	TCA 2403	
RESULT 3			
LOCUS	AY091075	2637 bp mRNA linear	PLN 18-SEP-2002
DEFINITION	Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.		
ACCESSION	AY091075		
VERSION	AY091075.1	GI:20259550	
KEYWORDS	FLU CDNA.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana (thale cress)		
REFERENCE			
AUTHORS	Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2637)		
AUTHORS	Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		

TITLE	JONES, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
JOURNAL	Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.
FEATURES	The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banb, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
source	Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.	
location/Qualifiers	1. .2637
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/chromosome="5"	
/clone="RAFL09-76-G11 (R19395)"	
/ecotype="Columbia"	
/note="This clone is in a modified pbluescript vector (pUC-1) as a BamHI/XhoI insert."	
1. .2637	
/gene="At5g42480"	
1. .113	
/gene="At5g42480"	
114. .2519	
/gene="At5g42480"	
/codon_start=1	
/evidence="experimental"	
/product="unknown protein"	
/protein_id="AA13895.1"	
/db_xref="GI:20259551"	
/translation="MRALSHVGTGSLSPOLCRLPATTKLRSHSTTTICSAKNADRLISDPNFTSDSSSPATRTATTTATVSPSPIDREPRNVPITDIYQUGAOTHTPLTGIRRAFEARVSKRPQGFSDDLISRRQLQAACSTLSPRSRRYNEGLDDEEXTVITDVPMDKVPFALCVLOSGGTEIYLRVGEALIKRLPSPSQDVLVVALAFVLDSDMAADPPDPFIFGYEFVEALGLDSEGSALPRLRAIDITLEIFRYPVLELLGLPGLDGYAARLNGSLGVRNLLMSVGGGASALVGLTTEKFNBAFLRWTAAEOYDLFVATPNTNIAKRSFEYEVVALVAQAFTGKKPHLLDDAKQFQQLQAQKMAEIPAMLVTRNNWEIDFLEERGLCALIGKYDECKMVLGDSBSQIRNPATIVERVLSNNDLNDLDPGLCKLETWLAGVFPFRFDTKKFKGLDGYDDPWLSTYLRREVVGSGPLAAAAAARACAEHVAKASAAQLOKVPFSPYTDNSAEPXOVQTFVSVPVGNVGDAGGCVFAEVRSENFENDYAIRAGVESVDETVMSVADMLKEASKTLIAGVGLGILSPFOKYFLKSSSFORDWVSNESVDATIGSVRADDSLEPRMDARTAEIVSKWQKIKSLARQPDRIEMLPVEDLGRMLKLTPTDAAITAOGLVYDDTLTLGLSDSVVVSAGCTALVEATLESACSLDLVHPENNADTVRTYTRIEVFSKGMKITEGSLAS"	
2520. .2637	
/gene="At5g42480"	
ORIGIN	
3'UTR	
Alignment Scores:	
Pred. No.:	1.68e-248
Score:	4051.00
Percent Similarity:	99.8%
Beet Local Similarity:	99.8%
Length:	2637
Matches:	799
Conservative:	0
Mismatches:	2

Query Match: 99.7% Indels: 0
DB: 15 Gaps: 0
US-10-600-070b-2 (1-801) x AY091075 (1-2637)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCybArgLeuPro 20
DB ATGGAAGCTCTGAGTCACTCGGCAATGCTCTCCCATTCCTCAATATATCCATTCACA 173

QY 21 ProAlaThrThrIleuLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
DB 174 CCGGCACGCAAGCAAGCTCCGACGAGCAGCAACACCTCTCAACTATCTGCTCCGACG 233

QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
DB 234 AATATGGGCGCAACGCTCTCTCTCGACTTCAATTTCCATTCCTCCATCTCTCTCTCC 233

QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
DB 294 TTCGCCACCGCACCAACCAACGCACTCGTCTCTCCGCCACCATCTATTGATCGTCCC 353

QY 81 GluArgHisValProIleProIleAspPheTyrGlnIleuGlyValArgIleThrHisPhe 100
DB 354 GAACGCCACGCTCCCATCCCATCTGATTTCTACCGAGTATTAGAGCTCAACACATTTCC 413

QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleYsProPheGlnPheGly 120
DB 414 TTAAACGATGGAATCAGAAAGCACTTCGAAAGCTTTCGAAACCGCGCAATTCGCT 473

QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleuGlnAlaCysGluThrLeu 140
DB 474 TTCAGGACGACGCTTAAATCAGCCGAGACAGATTCTTCAGAGCTGCTCCGAAACTCTG 533

QY 141 SerAsnProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspAspGluAlaThr 160
DB 534 TCTAAATCTCGGCTAGAAAGAGTACAAAGAGCTTCTTGAATGAAAGAAAGCTACA 593

QY 161 ValIleThrAspValProTrpAspValProGlyAlaLeuCysValLeuGlnGluGly 180
DB 594 GTCACTCATGATGTTCTTGAGATTAAGTCTCTGCTCTCTCTGTATGTCAGAAAGGT 653

QY 181 GlyIleuThrGluIleValLeuArgValGlyGluAlaLeuLeuYsGluArgLeuProYs 200
DB 654 GGTGAAGCTGAGATGATCTTCGGGCTGTGAGGCTCTGCTTAAAGAGAGTTCCTTAAG 713

QY 201 SerPheIleuGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
DB 714 TCGTTTAAAGCAAGATGTGTATTAGTATGCGCTTCGTTTCGATGTCTCGAGGGAT 773

QY 221 AlaMetAlaLeuAspProPheAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
DB 774 GCTATGACATTTGATCCACTGATTTTAACTGATTAAGCTTTGTTGAGAAAGCTTTG 833

QY 241 LysLeuLeuGlnGluGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
DB 834 AAGCTTTTACAGAGAGAGAGAGCAAGTACGCTTTCACCGGATTTACGTGCACAATTTGAT 893

QY 261 GluThrIleuGlnGluIleThrProArgTyrValIleuGluLeuLeuGlyLeuProLeuGly 280
DB 894 GAGACTTTGAGAGAGATCACTCCGCTTATGTCTTGGAGCTACCTTGCTTAACCGCTTGGT 953

QY 281 AspAspPheTrpAlaIleLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
DB 954 GATGATTAACCTGCGAAAGCACTAAATGTTTAAAGCGGTGTCGAAATTTTGTGTCTCT 1013

QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluYsPheMetAsn 320
DB 1014 GTTGGAGAGGTGAGATCACTCTTGTGGGGGTTTGAACCGTGAAGATTATGAT 1073

QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
DB 1074 GAGGCGTTTTTACGAATGACAGCTGCTGAGAGGTTGATCTTTTGTATGCTAACCCCAACG 1133

QY 341 AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAla 360
DB 1134 AATATTCAGCAGAGATCATTTGAAGTTTACGAAGTTCACCTTGTCTTGTGCTCAAGCT 1193

QY 361 PheIleGlyLysLysProHisIleLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
DB 1194 TTTATGTGTAAGAGCACACCTTTTACAGAGTCTGATGAGCAATTTCCAGCACTTCAG 1253

QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrp 400
DB 1254 CAGCTTAAGCTAATGCTATGAGATCTCTGCGATGTTTATGATACAGGAATATTTGG 1313

QY 401 GluIleAspPheGlyLeuGlnArgGlyLeuValAlaLeuLeuIleGlyLysValArgGlu 420
DB 1314 GAGATAGACTTGGCTTAAAGAGGAGACTGTGCACTGCTTATAGGCAAAAGTTGATGAA 1373

QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
DB 1374 TGCCTATGTGTGTGGGCTTAAACAGTGAAGATTCAATATAGGAATCCAGCTATTGTG 1433

QY 441 GluPheValLeuGluAsnSerAsnArgAspAspAspLeuProGlyLeuCysLys 460
DB 1434 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGGACTATCCAA 1493

QY 461 LeuLeuGluThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
DB 1494 TTGTTGAAAACCTGTTGGCAGGGGTTGCTTTCCTTCAAGGTTTCAGAGACACCAAAAGTAA 1553

QY 481 LysPheLysLeuGluLysAspTyrTyrAspAspProMetValLeuSerTyrLeuGluArgVal 500
DB 1554 AATTTTAACTCGGGACCTATGATGATCCATGATGTTTGAATCTTGGAAAGAGTG 1613

QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
DB 1614 GAGGTATCTCAGGGTCTCTCTTACGCTGCTCGACGATGCAAGATTGGAGCCGAG 1673

QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
DB 1674 CATGTAAAGCTATGCTATGACGACCTGCAGAAAGTTTTCCTCCGCTATACAGAT 1733

QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
DB 1734 AGAACTCGGCTGAAACCAAGATGTGCAAGACAGCTTTATGTTAATCTCTGTGGT 1793

QY 561 AsnAsnValGlyArgAspGlyLeuProGlyValPheIleAlaGluAlaValArgProSer 580
DB 1794 AACATGTAGGCGGTATGATGCTGCTGCTGCTTTTATGCAAGACTGTAGACCTCTCT 1853

QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
DB 1854 GAABACTTTGAACCTAATGATTTATGCAATTCGAGCTGGGGTCTCAGAGATAGCTTGAT 1913

QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu 620
DB 1914 GAACCTACCTGTGAATGTCCGTTGCTGATGTTTAAAGAGGCAAGTGAAGATCTTA 1973

QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleTyrPheLeuLysSer 640
DB 1974 GCTGCTGAGTGTGCAATTTGACCTGATTTCACTTACGCCAGAAAGATTTTCTTAAAGC 2033

QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
DB 2034 AGCTCATCTTTTCAAGCAAGATATGCTTTCTTATGGAATCTGATGTCGTACCATTA 2093

QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
DB 2094 GGCTCAGCTCAGACTCAGATTCAGTTCAGAGACCTTCCAGAAATGATCTCAGACTCAGAG 2153

QY 681 AsnIleValSerIleTyrGlnLysIleYsSerLeuAlaPheGlyProAspHisArgGly 700
DB 2154 AATATGATATCCAAAGTGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2213

QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTrpThrAspArgAlaAla 720

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	PIRMBED	REFERENCE	AUTHORS	FEATURES	source	gene	mrna	CDS
Db	2214	GAATCTTACCGAGGTTTGGATGGCGCAATGCTGAAGATTGGACTGACAGCAGCT	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Qy	721	GIJTHRALIGLNUGLYLeuValIYAspTYrThrleuLeuYleuSerValaspSer	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Db	2274	GAACCTGCAGCGCTGGGTGGTTTATGATTATATACCTGTGAACCTATCTGTTGACAGT	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Qy	741	ValITHrValSerIAspGlyThrArGAlaLeuValGIuValarThrLeuGIuGIuSerAla	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Db	2334	GGACAGCTTCAGCAGATGGAACCCGCTGCTGGTGGAACCACTCGAGAGAGCTGCT	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Qy	761	CybleuSerAspLeuValHisProGIuAspAspAlaThrAspValArGTYrTYrThr	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Db	2394	TGTCATCTGATTGGTTGTTTCATCCAGAAACAAAGCTACTAGTCAAGACCTACACACA	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Qy	781	ArgTYrGIuValPheTYrSerIYsSerGIYrTYrLeuIleThrGIuGIuSerValLeuAla	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Db	2454	ACATACGAAGTTTCTTCGCTCAAGTACAGGCTGGAAATCACTGAAGCTCTGTTCTTGCA	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Qy	801	Ser 801	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS
Db	2514	TCA 2516	AY221468	1	GI:33436338	Arabidopsis thaliana (chale cress)	Arabidopsis thaliana	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary	Vitina, S., Kosharova, O., van Erp, H., Froehlich, J.E., and	Direct Submission	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State	12897262	1	3668	Location/Qualifiers	1. 3668	gene	mrna	CDS

[illegible]

QY 213 aPhelLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyr 233
 Db 1198 GTTTCGATGCTCTCGAGGATGCTATGCATTCGATCCACCTGATTTATTTACTGCTTA 1257
 QY 233 rGluPheValGluGluAlaLeuIleLeuGln----- 244
 Db 1258 TGAGTTGTTGAGGAGCTTTGAAGCTTTTACAGGTAGTTGACTTCTTGGAATTTTG 1317
 QY 244 ----- 244
 Db 1318 ACGAGCTTGCGCTTATATAGAACTTTCTGATTTGATTACTTTGATGCTTGCTGA 1377
 QY 245 -GluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGln 264
 Db 1378 GGAAGAGAGGCAAGTAGCTTGCCGACCGGATTTACGTGCACAAATTGATGAGCTTTGCA 1437
 QY 264 uGluIleThrProArgTyrValLeuGluLeuGluGlyLeuProLeuGlyAspAspTyrAl 284
 Db 1438 AGAGATCACTCCGCGTATGCTTGGAGCTACTTGCGCTTACCGCTTGATGATTTACGC 1497
 QY 284 aAlaIleArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuThrSerValGlyGly 304
 Db 1498 TGGAAAAGAATAATGCTTAAAGCGGTGTCGGAATATTTTGCTGTGGAGGAGG 1557
 QY 304 yGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluIlePheMetAsnGluAlaPheLe 324
 Db 1558 TGGAGCATCAGCTCTTGTTGGGGTTGACCCCGTAGAAGTTATGATGAGCGCTTTT 1617
 QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
 Db 1618 ACGAATGACAGCTGCTGAGCAGGT-ATACAGTTTATGATACCTTTTAAATTTCTTTAGC 1676
 QY 333 -----AspLeuPheVal 336
 Db 1677 ATGATATTAACCTTAGGTTCTCATTTTAATGATATGTTGTGTGAGGTGATCTTTTGT 1736
 QY 336 lAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLe 356
 Db 1737 AGCTTACCCCAAGCAATTTCCAGCAGAGTCACTTTGAAGTTTACGAATGTGACTGCTCT 1796
 QY 356 uValAlaGlnAlaPheIleGlyIleGlyIleProIleLeuGlnAspAlaAspIleGlnP 376
 Db 1797 TGTGGCTCAAGCTTTTATGTTGTAAGAAGCCACACTTTTACAGGATGCTGATTAAGCAAT 1856
 QY 376 eGlnGlnLeuGlnGlnAlaIleValMetAlaMetGluIleProAlaMetLeuTyrAspThr 396
 Db 1857 CCAGCAACTTCAGCAGGCTAAGTAACTGATGAGATTCTGCGATGTTGATGATAC 1916
 QY 396 rArgAsnAsnTrpGluIleAspPheGlyLeuGluArgGlyLeuGlyAlaLeuLeuIleG 416
 Db 1917 ACGGAATATTTGGAGATGAATCGCTTCAGAAAGGAGACTCTGTGCACTGCTTAAGG 1976
 QY 416 yIleValAspGluCysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAs 436
 Db 1977 CAAGTGAATGAATGCCGTATGCTTGCGGCTTACAGTGAAGGATTCACAAATATAGAA 2036
 QY 436 nProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAsnAspAspLeuP 456
 Db 2037 TCCAGCAATTTGAGAGTTGTTTGGAGAAATTCAAATCGATGATCAATATGATCTCCC 2096
 QY 456 oGlyLeuGlyValLeuLeuGlnIleThrTrpLeuAlaGlyValValaPheProArgPheAspAs 476
 Db 2097 TGGACTATGCAAAATGTTGGAACCTGTTGGCAGGGGTTGCTTTCTTCAAGTTCAAGA 2156
 QY 476 pThrIleAspIleValSerPheIleLeuGlyAspTyrTyrAspAspProMetValLeuSerTyr 496
 Db 2157 CACCAAGATTAATAAATTTAACTCGGGAATACTATGATGATCTATGCTTTGACTTA 2216
 QY 496 rLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArg 516
 Db 2217 CTGGAAAGAGTGAGGTAGTTCAAGGGTCTCTTATGCTGCTGCAACTATGAGCAG 2276

QY 516 gIleGlyAlaGluHisValIleValAspAlaMetGlnAlaLeuGlnIleValPheProSe 536
 Db 2277 GATTGAGCCGAGCAGATGATAAGCTATGCTATGACGAGCACTGCAGAAAGTTTCTTCTTC 2336
 QY 536 rArgTyrThrAspArgAsnSerAlaGluProIleAspValGlnGluThrValPheSerVal 556
 Db 2337 CCGCTATACAGATGAAACTCGGCTGAAACCAAGATGTCAGAAAGACAGCTTTTAGTGT 2396
 QY 556 lAspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValaPheIleAlaGluAl 576
 Db 2397 AGATCTGTTGTTGTAACAATGTAGGCGGTATGCTGAGCCGTCTTTATTTACAGAGC 2456
 QY 576 aValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerG 596
 Db 2457 TGTAAAGCCTCTGAATAAATTTGAAACTATGATTTGCAATTCAGCTGGGCTTCAGA 2516
 QY 596 uSerSerValAspGluThrThrValGluMetSerValaIleAspMetLeuGluIleAsp 616
 Db 2517 GAGTACCGGTGAAGAACTATGTTGAATGTCCTGTTGCTGATATGTTAAAGAGCGCAG 2576
 QY 616 rValIleIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleTyr 636
 Db 2577 TGTGAAGATCTTACGCTGCTGCTGCTGCTGCAATTCGACTGATTCATCTTACGCCAGAAGTA 2636
 QY 636 rPheLeuIleSerSerSerSerPheGlnArgIleAspMetValSerSerMetGluSerAs 656
 Db 2637 TTTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTCTTCTATGGAATCTGA 2696
 QY 656 pValAlaThrIle----- 660
 Db 2697 TGTGCTACCAATGATATGATTAATGATGCAATTTTCATATATCTGCAATTCCTCAAAT 2756
 QY 660 ----- 660
 Db 2757 ATGCTGTTTGTGACCTAAGAACATATGTTCCACTTAATACATGTCCAAAGTTGTAC 2816
 QY 660 ----- 660
 Db 2817 CAGATTAACAAGTTGCTGATGATTAATTTCACTAATATGCTGTTGAATTTTGTGATCA 2876
 QY 660 ----- 660
 Db 2877 ACTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATACGTAGATTA 2936
 QY 660 ----- 660
 Db 2937 GAGATTCCTTAGTGCGCTTGTCCAACTTTTCTTCTGATTTTCTTTTGATTT 2996
 QY 661 --GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
 Db 2997 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTAGAGACTGAG 3056
 QY 680 lAsnIleValSerIleTrpGlnIleIleIleSerLeuAlaPheGlyProAsnHisArg 700
 Db 3057 AGAATATAGTATTCAGATGCGACAGAAATTAAGTCTGCTTTTGGGCTGATCACCAGCA 3116
 QY 700 lGluMetLeuProGlu----- 705
 Db 3117 TAGAAATGTTACAGAGGAGGAGAAATTAATCTTAATCAATTCATCAATTTGTTGAAAACTG 3176
 QY 706 -----ValLeuAspGly 709
 Db 3177 TTGAGCATGATTATAGTCGTGCTGTTGATTCGTGTTATTTATGATTTGATGAGG 3236
 QY 710 ArgMetLeuIleIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyr 729
 Db 3237 CGAATGCTGAAGATTTTGACTGACAGAGCAGCTGAAACTGCGAGCTTGCTTGCTTAT 3296
 QY 730 AspTyrThrLeuLeuIleSerValAspSerValIleAspSerValThrValSerAlaAspGlyThrArg 749
 Db 3297 GATTATACACTGTTGAATCTATCTGTTGACAGGTGTACAGTCTCAGCAGATGGAACCCGT 3356
 QY 750 AlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGlu 769

```

|||||
Db      3357 GCCTGGTAGAGCACTCGAGAGCTGCTGTCTATCTGATTTGGTTCATCCAGAA 3416
|||||
Qy      770 AAsaAaLaTharPpAlaGrThrThrThrArgTyrgValPheTIPSerLySer 789
|||||
Db      3417 AACAAAGCTACTGATGTCAGAACTTACACACAGATACGAAAGTTTCTGCTCCAGTCA 3476
|||||
Qy      790 G|YrPlySilerThrgLugLySerValIleuAlaSer 801
|||||
Db      3477 GGGTGGAAATCACTGAGGCTCTGTTCTTGATCA 3512
|||||

RESULT 5
AB016888      85791 bp  DNA  linear  PLN 14-FEB-2004
LOCUS         Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDH9.
DEFINITION   AB016888 BA000015
ACCESSION    AB016888.1 GI:3449329
VERSION      AB016888.1 GI:3449329
KEYWORDS
SOURCE
ORGANISM     Arabidopsis thaliana (thale cress)
              Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS      1 Asamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kotani, H., Miyajima, N.
              and Tabata, S.
              Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
              Sequence features of the regions of 1,081,958 bp covered by
              seventeen physically assigned P1 and TAC clones
              DNA Res. 5 (6), 379-391 (1998)
JOURNAL
PUBMED
AUTHORS      2 (bases 1 to 85791)
              Nakamura, Y.
              Direct Submission
              Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research
              Institute, Department of Plant Gene Research, 1532-3, Yana,
              Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
              Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT
              Address for correspondence: kaos@kazusa.or.jp
              For the latest information on annotation of this clone, please see
              http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDH9
              Genes with similarity to proteins in the databases are described in
              'product' or 'note' qualifiers. Genes that have no significant
              protein similarity are described as 'unknown protein'.
              The software programs used to predict genes include: Graal
              (Informatics Group, Oak Ridge National Laboratory,
              http://compbio.ornl.gov/Graal-1.3/),
              GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
              NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
              Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
              SplicePredictor (Volker Brendel, Stanford University,
              http://gremli1.zool.iastate.edu/cgi-bin/sp.cgi).
              Genes encoding tRNAs are predicted by tRNAscan-SE
              (Sean Eddy, Washington University School of Medicine, St. Louis,
              http://genome.wustl.edu/eddy/tRNAscan-SE/).
              This sequence may not be the entire insert of this clone. It may be
              shorter because we remove overlaps between neighboring submissions.
              The 5' clone is K5J14 and the 3' clone is K16E1.
              Location/Qualifiers
                1. 85791
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
                   /chromosome="5"
                   /clone="MDH9"
                   /clone_1lb="Miteui P1"
                   /ecolyse="Columbia"
                   /complement(59..239)
                   /note="CDS is reported in Acc# AB023032
                   gene_id:K5J14.12
                   unknown protein"
                   /number=1
                   /evidence=not_experimental

CDS
complement(join(2056..2079,2278..2293,2991..3633))
/note="contains similarity to unknown protein"
gene_id:MDH9.1
protein_id:BA010474.1
|104864
|pseudo
/codon_start=1
/evidence=not_experimental
join(5739..6138,6462..6595)
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BA010474.1"
/db_xref="GI:9759469"
/translation="MKYDEPAEEDRKKSSNLISLAKNSVPLPERGRFYREYSKRRE
RLRRKGEEAEGVVGKTPYNLGVNHEPTNKRKGAKESIKTKVSNVSSSTIR
YSLRMAKENRKPPLPLNVAVSAMKSVYTRKEBRLSLICTACTNTSSGLLAFPLTGS
EPLGVHKRTLPSPALV"
complement(join(6868..7911,7989..8414,8555..8671,
8754..9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not_experimental
/product="am repeat containing protein"
/protein_id="BA010475.1"
/db_xref="GI:9759470"
/translation="WTDDEBATGDAENRDEEVTAEPIHDEYVDVAEIHSEVKEDD
DCGLVSDIVSVEFLDQINGVRRDQCEFNVRRLKILPLADEIRGESPCKHP
LNRIRKVFPLAKKLELTGNSGSIYALDETMTRFHSIYKRLNVLVAPDELM1
SGDAKDEIDSLCKOLKAKRRTPDIDELAVMVAVSSTKDPNADALIERLAKLE
LQIDDLKTEITIAIOSLIDPKGLNIEETKOH1IELNKKGLGAEADIIYGVNIK
AIRKSTSLIPHEFLCPILEIMLDPVITATGQYREKESIQKMPDAGTKCPTRSL
DHLSLAPNFKRLIMQCEKKNPKIPEKVSPOQEQKQVSLVLAELSSQLEQ
RSVKQMRLLARENPENRVLINAGAIPLVQLSLYPDSIQENAVTTLNLIDEVL
KKLISNEGAIPINIEILENGNREARENSAALFSLMLDENKVTIGLSNGIPLVLDL
QHTLGRKQDALFALFNLISLNSKRAIDAGIOPVLNLIKOKNLGMDIALSLIL
LASHPEGRQAIQGLSFLETLVEFIRGQTPNKRCASVLELSNNSSTFLAALQFV
YEVLTITSGTRAKRKANALIGLSKEQT"
10888..12579
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BA010476.1"
/db_xref="GI:9759471"
/translation="MMISKPLVEESIRDLLEVLTVSRRLVYSVQKLKKKIKHTEV
VEDEIARAVNCLISVGCRAVDTEDEDSNKKMSASEGKGLMTICGEBETL
DCSSVGRERRPKKNNRKYLDAGDGRKHVYLPDILMCLMRLLPLTSLNHLVCK
KMSMANQRFLOMRREGSPQTPMLPLPALADGCGSGDIGHGVQODKRRH1ETDL
KGRFMSVTSIHETIYVGRSMDRNSFSSHGILVPSITKMRKTIASRHAASLPT
VGAETVSESTQOTKONRODRFHLRSVGEEDVVEDPRLRSVRQNRNSAQONGTK
SHRLIROKDLRLNRNSKRFVLIAGTGLFDEPLDSGEIYDASTWMSQMLPMGF
GVVSGGILNGIFAYASNDKLSGYIIEGFWLTITQSP1PPEVHEFYPGLVCSNRL
FMLSVCMBQEGQIGRRKAVKAMELDLVLTWTEVSHNPAADMAATVYSDDNI
LMSIENFKIFGQYLSFEFTYCDVUTEASNRHRSRNRQSLNLSCTNKTIALHL"
complement(15087..16778)
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BA010477.1"
/db_xref="GI:9759472"
/translation="MMISKPLVEESIRDLLEVLTVSRRLVYSVQKLKKKIKHTEV
EDBEVARGAVNCLISVGCRAVDTEDEDSNKKMSASEGKGLMTICGEBETL
DCSSVGRERRPKKNNRKYLDAGDGRKHVYLPDILMCLMRLLPLTSLNHLVCK
KMSMANQRFLOMRREGSPQTPMLPLPALADGCGSGDIGHGVQODKRRH1ETDL
KGRFMSVTSIHETIYVGRSMDRNSFSSHGILVPSITKMRKTIASRHAASLPT
VGAETVSESTQOTKONRODRFHLRSVGEEDVVEDPRLRSVRQNRNSAQONGTK
SHRLIROKDLRLNRNSKRFVLIAGTGLFDEPLDSGEIYDASTWMSQMLPMGF
GVVSGGILNGIFAYASNDKLSGYIIEGFWLTITQSP1PPEVHEFYPGLVCSNRL
FMLSVCMBQEGQIGRRKAVKAMELDLVLTWTEVSHNPAADMAATVYSDDNI
LMSIENFKIFGQYLSFEFTYCDVUTEASNRHRSRNRQSLNLSCTNKTIALHL"

```

exon

CDS

LMGVEMFKIEGVGLSPTVCDILTEBSRWHSRNOROKNLSCNKTIALIHL"
complement(join(18111..18193,18281..18535,18773..19001,
13088..19195,19307..19447,19542..19952,20129..20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
epiP42251"

CDS

/codon_start=1
/evidence=not experimental
/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFWCLSLTLTVIAEETPTIRIAFGSCANQAPQIWMDA
INKFDPOLFTWLDNDIYGDIRPLTVIGKERTGPWRNSPRFVPSSEEMKRYAAKQ
ANPGYSIORKAKVIGTMDHDGGLNDAGKEPRKYNQKMLDPLNEPLDSRRQA
GYASATYGPNNKRVIVLDPYHNDPLRSQSLIGDQMDLMEALSGPSSEITII
GSSVOYISLNSATGPFYFMSWGRPKERKIFKILADTKRQGVFISGDVHREGIT
RDQSVGYPLVDYSSGLVQSVKVPRLSRIVRLFWTPSTMRVINDCKFSCT
YGOONFGAISIDNNANPVTIKLEIRDQNGVIGTVNSLIGAVLAPGSNSLKDOTTGK
SORCYTEIEIPTRRYLVIYFTIAYLMAAILGLIIGAVLAIRACYKCKVD"
complement(21540..22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"

CDS

/translation="WTLAKNOKSSLRLYKYSKRSSESRNIEDSRSSNSGSSS
LVNELRTVFEDYDANSDKISGELOSQCSVLGLMSSEVEVKTSDDVDGID
FEFFLKLMEGSDSDESRKELKEAFMYMEGEERTYASLRRTLSRIGESCITYAC
KVMIRGFDDNDVLSFDEFLMMR"
join(24090..24289,24410..24614,24701..24945,25115..25313,
25632..25679,25821..25928,26039..26227,26314..26613,
26730..26831,27169..27303,27487..27711,27919..28122,
28213..28380,28646..28795,28889..29026,29204..29431,
29519..29671,29784..29910,30115..30241,30554..30674,
30763..30843,30932..31033,31112..31262,31338..31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="p1trf1lyb1n"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFTVGVKFSPIILAPNSGDSRRGRVYLKDSRNKYRPNPS
SPRLTPRRARVAEASLIPYNGLMAOPNSHGKGLKNITVSGKATGSLISQGNPCLT
CKNQGIRRALPSAFVDRYAFSLRSLSLSLRKSQIYNATLGDPEHPAATAMPD
GIYAEKQDLDLPEEIDSAELAEALFGLCELPSPKLRGQKNGRLRYILPNKVPNRK
EAMVEYVSGSIDEEDEQCIAMHIEHVAFGSKRREKLGCTGRSNAAYDPPHIVPHI
HSPHTKDSDDLFPSPVLDALNEIAFHPKFLSRVERKRAIISLQMMNTIEKYDC
OLLOHASENKIGRRPIGLEBOIKKMDVKIRKPHRMVFPANATLYIGDIDNIPR
TYNITEAVFGKGLDNSTSSPSPPAFGMANFVYKIPAGICGTPSNKDTQATG
KMKIREHAIKRPVEHNWSLPGTSVDLKPPQIFRHELLQNPALIMPKFIPSVKQTG
DLNVLNMRKIFLSALHFRINTRYKSNPPTSVELHSDSGRGCTVTLTVAEPPN
WONAKVAVQEVRLKEFGVTRGELTYMDALKDSHLLAMIDNVSVDNLVIMES
DALSHVTMDQTOGHETLVAVAGVTLLEVTYVGAKEVLEFISDFGRPYAPALVACV
PTRYNDGSGSDPNTSPDELTISVKSGLAPIEABEIPSEKELISQGLKELLTOR
NCPFVPIPGSGKTKLHDKETGTQRLSNKIANYKSTTESAPAGWELVGGGRABE
TSDSKAIVVGVRLTSEGRVGPDSREQVELFCVNLINCSLSTEFIMMERFTLR
DNQMOAFQILHVLNERSVWLEDAFDAROLDYLSVFSFSEELERAAHMLAMNG
DERFEVETPKSLQSLNLSYDAVMHSEHEDNMEVSIVGFSSEELERAAHMLAMNG
ASHDSKAPGSEPIIPROPTAGLOFOVPLKADIERCAIAGAPARMFOVTDGDL
FOSVSKLPVADHGLIKSEBOLLFGGBDELOKRLRHPALFGVIMGLAELIINRLFT
VROSLGLTVYVSFEALNFDRLNIGVYISTSTPGKYKAVADKSVLRLSHNOTAP
RELDRAKRLILMRHEELKSNATWLNLAQSSVREKLSCKGVVLYEASIID
IYLAYNQLRVDEBDSLYSICIGAGQAEELITVLSSEBEPRKLSCKGVVPGVGRGSSMTTR
PTT"

CDS

complement(join(33251..33326,33512..33663,33744..33797,
33868..33968,34143..34265,34359..34768,34873..35052,
35143..35265,35387..35509,35631..35753,35875..36000))

Alignment Scores:

Pred. No.: 2.87e-232 Length: 85791
Score: 3825.50 Matches: 798
Percent Similarity: 78.7% Conservative: 0
Beet Local Similarity: 78.7% Mismatches: 3

Query Match:	94.2%	Indels:	214
DB:	15	Gaps:	5
US-10-600-070b-2 (1-801) x AB016688 (1-85791)			
Qy	1	MetGluAlaLeuSerThiValJylIegIleuSerProPhcGluLeuCyArgLeuPro	20
Db	64077	ATGGAAGCTCTGAGTCACGTCGGCATTTGGTCTCCCAATTCGAATTAACGATTAACA	64136
Qy	21	ProAlaThrThyLeuArgArgSerHisAanThrSerThiIleCySerAlaSer	40
Db	64137	CCGGCAGCACAAGCTCCGAGTGACCAACACCTTCAACATCTGCTCCGCAAC	64196
Qy	41	LySTPAlaAspArgLeuSerSerAspPheAanPheThrSerAspSerSerSer	60
Db	64197	AAATGGCCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCC	64256
Qy	61	PheAlaThrAlaThrThrThraAlaThrValSerLeuProProSerIleAspAlaPro	80
Db	64257	TTGCGCACCGCACCAACCGCCACTCTGCTCTCCGCCACATCTTAATTCCTCC	64316
Qy	81	GluArgHisValProIleProIleAspPheTyArgValIleuGlyValGlnThrHisPhe	100
Db	64317	GAAACGACGTCCTCCATCCCATGATGTTCTACCAAGTATAGAACTCAACACATTC	64376
Qy	101	LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleProGlnPheGly	120
Db	64377	TTACCGATCGAATCGAAGAGATTCGAAGCTAGGCTTCCGAACCGCGCAATTCGG	64436
Qy	121	PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCyGluThrLeu	140
Db	64437	TTACGAGCAGACGCTTAATTCAGCCGAGACAGATTCCTCAAGCTGCTTCCGAACCTG	64496
Qy	141	SerAspProArgSerArgArgGlyIleArgGluGlyLeuLeuAspAspGluGluAlaThr	160
Db	64497	TCTAATCCTCGCTGTAAGAGAGTACATGAAAGCTTCTTGATGATGAAGACTTAC	64556
Qy	161	ValIleThrAspValProThrAsp	168
Db	64557	GTCATACGATGATCTCTTGGATAGTAATTCGATTTCGGAATTAAGATTCTTCTC	64616
Qy	169	-----LyValProGlyAla	173
Db	64617	GTTTAATTCATGAATGGAATGAAGAAAGAACTTTTACTGATGAAGGTTCTCTGTC	64676
Qy	173	AlaLeuCyValIleuGlnGluGlyGlyIleuThrGluIleValIleuArgValGlyGluAlaLeu	193
Db	64677	TCTCTGTGATTCGAAAGAGTGTGAGACTGATGATGTTCTCGGATGGTGAAGCTCT	64736
Qy	193	IleuLeuGluArgLeuProIleSerPheIleGlnAspValIleuValIleuAlaLeuAla	213
Db	64737	GCTTAAGAGAGGTTGCTTAAGCTGTTTAAGCAAGATGCGTTTATGTTATGCGGCTGC	64796
Qy	213	AspLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTy	233
Db	64797	GTTTCGATGCTCGAGGAGTCTATGCACTGATTCACCTGATTTAATACGATTA	64856
Qy	233	GlyIlePheValGluGluAlaLeuIleLeuGln	244
Db	64857	TGAGTGTGTAAGAAAGCTTTGAAGCTTTACAGTATGATTCGCTTGGTAAATTGG	64916
Qy	244	-----	244
Db	64917	ACAGCGCTTGGCTTTAAGAACTTCTTGATTTGATCTTGTATTTGAGCTTGTGTA	64976
Qy	245	-GluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGly	264
Db	64977	CGAGGAAGAGCAAGTAGCTTGCACCGATTAACGACCAATTCATGATGAGACTTTGGA	65036
Qy	264	uGluIleThrProArgTyValIleuGluLeuLeuGlyIleuProLeuGlyIleAspAspTyAla	284
Db	65037	AGAGATCACTCCGCTTATGTTGAGGACTTACCTTGGCTTACCGCTTGGATGATTAATGCG	65096

QY 284 aalalyarxleuansgilyleusergilyValArgasnilleuTrpserValglyylgi 304
 DB 65097 TGGGAAAAGCTAAATGCTTTAAAGCGGTGCGCGAAATTTTGGGTCTGTGGAGGAGG 65156
 QY 304 yglYalaserAlaleuValglyleuThraArgLylapheMeransgilyalphele 324
 DB 65157 TGGAGCATCAGCTCTTGTGGGGGTTGACCCCGAGAGAGTTTATGAATGAGGGGCTTTT 65216
 QY 324 uArgMeThraAlaleuVal----- 332
 DB 65217 ACCAAATGACAGCTGTGAGCAGGT-ATACAGTTTATACCTTTTAAATTTCTTTAGC 65275
 QY 333 -----AapleuPheVa 336
 DB 65276 ATGATATACCTTAAGCTTCTCATTTTAATGATATGTTGTGTGAGGTGATCTTTTGT 65335
 QY 336 lAlarhProSerAsnilleProAlagIuserPhegluValTyrGluValAlaleuAlale 356
 DB 65336 AGCTACCCCAAGCAATATTCACAGCAGATCATTTGAAGTTTACGAAGTTGCACCTGTCT 65395
 QY 356 uValAlagInAlapheIleGlylylylyProhileuLeuGluAaplyrsglnh 376
 DB 65396 TGTGGCTCAAGCTTTTATTTGTGAAGCAACCTTTTACAGGATCTGATTAAGCAATT 65455
 QY 376 eglInleuGlnGlnAlaleuValMetAlameGluilleProAlameLeuTyrAapTh 396
 DB 65456 CCAACCACTTCAGCAGGCTTAAGCTTAAGCTTAAGCTTAAGCTTCGCAATGTRATGATAC 65515
 QY 396 rArgAsnAsnTrpGluilleaPheglileuGluArgLyleuCyAlaleuLeuilegi 416
 DB 65516 ACCGAATTAATTTGGAGATAGACTTCGGTCTAGAAAGGAGACTGTCGACCTGTTATAGG 65575
 QY 416 ylylyValaPergLycArgMeTrpLeuGlyleuAapSerGluAapSerGlnTyrAag 436
 DB 65576 CAAAGTGAATGAAATGCGTATGTGTGGCTTAGACAGTGAAGATTCACAAATATAGAA 65635
 QY 436 nProAlaleuValGluPheValleuGluAanSerAsnArgAapAsnAapAapLeuPr 456
 DB 65636 TCCAGCTATTTGAGATTTTGTGTAATTCGTAATTCGTAATTCGTAATTCGTAATTC 65695
 QY 456 ogIlyleuCyAlaleuLeuGluThraTrpLeuAlaglyValValPheProArgPheArg 476
 DB 65696 TGAACATGCAAAATTTGGAAACCTGTTGGAGGGGTTGTCTTCTTAAGTTCAAGAA 65755
 QY 476 pThrlyAapLylapPheleuGlyAapTyrTyrAapAapProMeValleuSerly 496
 DB 65756 CACCAAAAGATTAATAATTTAACTCGGGACTACTATGATGATCTTATGCTTTGAGTTA 65815
 QY 496 rleuGluArgValGluValValGlnGlySerProleuAlaleuAlaleuAlaleuAla 516
 DB 65816 CTGGAAAGAGTGAAGTACTGAGGTCTCTCTTAAGCTGTGCTGCACTATGAGCAAG 65875
 QY 516 gIleGlyAlagLynhIleVallysaIaserAlameGlnAlaleuGlnlyValPheProse 536
 DB 65876 GATTGAGCCGACAGTGTGAAGTACTGTATGCAAGCACTGAGAAATTTTCTTCTTC 65935
 QY 536 rArgTyrTrhAapArgAsnSerAlagLuprolyAapValGlnGluThraValPheSerVa 556
 DB 65936 CCCCTATACAGATGAAGAACTCGGCTGAACCAAGATGTCAAGAGACACTGTTTACTGT 65995
 QY 556 lAapProValGlyAanAsnValGlyArgAapGlyGluProGlyValPheIleAlagLyn 576
 DB 65996 AGATCCCTGTGTGAACATATAGCCGTGATGGTGAACCTGTGTCTTTATTTGAGAGAC 66055
 QY 576 aValArgProSerGluAanPhegluThraAanApyrAlaleuArgAlaglyValSerGI 596
 DB 66056 TGTAAAGCCTCTGAAGAACTTTGAACATTAATGATTAATGCAATTCGACCTGGGCTCAGA 66115
 QY 596 uSerSerValaPergLuthraThraValGluMetSerValAlaAapMeLeuLyleuAlase 616
 DB 66116 GACTAGCGTGTGAAGAACTACTGTTGAATGTCCGTTGCTGATATGTTAAAGAGGCAAG 66175
 QY 616 rVallyrIleleuAlaleuAlaglyValAlaleuGlyleuIleSerleuPheSerGlnly 636

DB 66176 TGTGAAGATCTAGCTGCTGTGTGGCAATTGACGATTTCACTGTTGAGCCAGAACTA 66235
 QY 636 rPheLeuLysSerSerSerSerPheGlnaglyAapMeValIleSerSerMeGluSerA 656
 DB 66236 TTTTCTTAAAGAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGA 66295
 QY 656 ValAlaThrIle----- 660
 DB 66296 TGTGCTACCATAGATGATTAATAATGATGCAATTTTCATATATTCGATTCCTCAAA 66355
 QY 660 ----- 660
 DB 66356 ATGCTGTTTTGAGCTAAGAACATAGTCCCACTTAATACATGTCCAAAGTTGTAC 66415
 QY 660 ----- 660
 DB 66416 CAAATTAACAGTGTGAGTAATAATTTCATTAATGCTGCTGAATTTTGTATCAA 66475
 QY 660 ----- 660
 DB 66476 ACTGTAGACAGAAATGTAATTTCACTCAACATTTCTGTTAAGATTAACGTAGATTA 66535
 QY 660 ----- 660
 DB 66536 GAGATTCCTTAGTGTGCTGCTTGTCCAACTTTCTTCTGATTTTCTTTTGATTT 66595
 QY 661 --GlySerValArgAlaAapAapSerGluAlaleuProArgMeAapAlaArgThraIag 680
 DB 66596 TAGGTCACAGTCAAGCTGACGATTCAGAAAGCACCTCCAGAAATGATGCTAGACGTCAG 66655
 QY 680 luanlIleValSerlyrTrpGlnlyIlelySerleuAlapheglProAsnIleArgI 700
 DB 66656 AGAATATGATCAAGATGAGCAAGATTAAGTCTGCTTTTGGCTGATCACCGCA 66715
 QY 700 leGluMeLeuProGlu----- 705
 DB 66716 TAGAATGTTACAGAGGTGAGGAATTAATTCATCAATTCAATCAATTGCTGAAAACTG 66775
 QY 706 -----ValleuAapGly 709
 DB 66776 TTGACATGATTAATGATCTGTGCTGCTGTTGATGCTGTTATTAATGATTTTGGAGTGG 66835
 QY 710 ArgMeLeuLysIleTrpThraAapArgAlaleuGluThraAlagInleuGlyleuValTyr 729
 DB 66836 CGAATGCTGAAGATTTGAGACTGACAGACGCTGAACCTGCGAGCTGTGGTGTGTAT 66895
 QY 730 AapTyrTrhLeuLeuLysSerValAapSerValThraValSerAlaAapGlyThraArg 749
 DB 66896 GATTATACACTGTTGAACATATCTGTGACAGTGTGACACTTCACAGCAATGGAACCCGT 66955
 QY 750 AlaleuValGluAlaThreleuGlnGluSerAlaCyaleuSerAapleuValIleProGlu 769
 DB 66956 GCTCTGTGGAAGCAACTCGAGAGAGTCTGCTGTCTATCTGATTTGTTCTTCCAGAA 67015
 QY 770 AasnAlaThraAapValArgThryrTrhThraArgTyrGluValPheTrpSerlySer 789
 DB 67016 AACAAATGCTATGATGTCACAACTACACAAAGATGATTTTGTGTCACAGTCA 67075
 QY 790 GlyTrpLysIleThraGlnGlySerValleuAlaSer 801
 DB 67076 GGGTGAAGAAATCACTGAAGCTCTGTTCTTGATCA 67111
 RESULT 6
 AY221467 3664 bp DNA linear PLN 05-AUG-2003
 LOCUS Arabidopsis thaliana truncated division protein (arce6) gene, arce-1
 DEFINITION allele, complete cds, nuclear gene for chloroplast product.
 ACCESSION AY221467
 VERSION AY221467.1 GI:33436274
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3664)
Vilma, S., Froehlich, J. E., Kosharova, O., Pyke, K. A., Van Erp, H. and Otey, K. W.

TITLE
ARC6 is a 3-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein FtsZ
JOURNAL
Plant Cell 15 (8), 1918-1933 (2003)
PUBMED
12897262

REFERENCE
2 (bases 1 to 3664)
Vilma, S., Kosharova, O., van Erp, H., Froehlich, J. E. and Otey, K. W.

TITLE
Direct Submission
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
Location/Qualifiers

FEATURES

source
1. 3664
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Ws-2"
/note="genotype: arc6 mutant"
1. 3664
/gene="arc6"
/allele="arc6-1"
/join(478..984,1067..1291,1379..>1621)
/gene="arc6"
/allele="arc6-1"
/product="truncated division protein"
/join(478..984,1067..1291,1379..1621)
/gene="arc6"
/allele="arc6-1"
/note="truncation caused by premature stop codon"
/codon_start=1
/product="truncated division protein"
/protein_id="AAQ18644.1"
/db_xref="GI:33436275"
/translation="MRLSHVIGISLPOLCRLPATYKLRSHNTSTTCSAKMADRLSDFNFTSDSSSPATRTTATVSLSPSIDRPRHPRIPIIDYQVAGATHPLT
DGIKRAFEARVSRPQFGSDALISRLQIAACETLSNPSRRSRNEGLDDEBAT
VITVDWPKVGVLCVLOEGDETEIVLRVEALIKRLPKSFQDVVLWALFLVVS
RDMLADPPIRTGFEVERALKILQEGSSLAIDRAQIDITLEIETPRVYELILG
LPIDDDYAAKRLNGLSGVNRIILMSVGGGASALVGLTRKKFNNAEFL"

CDS

ORIGIN
Alignment Scores:
Pred. No.: 7, 72e-234 Length: 3664
Score: 3823.50 Matches: 798
Percent Similarity: 78.7% Conservative: 0
Best Local Similarity: 78.7% Mismatches: 3
Query Match: 94.1% Indels: 214
DB: 15 Gaps: 5

US-10-600-070b-2 (1-801) x AY221467 (1-3664)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuGlyArgLeuPro 20
DB 478 ATGAAGCTCTGATGTCAGTCGCGCATGGTCTCTCCCATTCATTAATATCCGATTACCA 537
QY 21 ProAlaThrThrIleValLeuArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
DB 538 CCGGCGAGCAAGCTCCGACGTAGCCACAACCTCTTCAACTATCTGCTCGCCAGC 597
QY 41 LysIlePheAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
DB 598 AAATGGGCGACGCTCTCTCTCGCACTTCAATTCACCTCCGATTCTCTCTCTCTCC 657
QY 61 PheAlaThrAlaThrThrAlaThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
DB 658 TTGCGCCACCGCCACACACCGCACCTCTGCTCTCTGACACCATTAATGATCGCC 717
QY 81 GluArgHisValProIleProIleAspPheThrGlnValLeuGlyAlaGlnThrHisPhe 100

DB 718 GAACGCGACGTCCCAATCCCATTTGATTTTCACAGATTAATGAGACTCAACACATTTCC 777
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleSerProGlnPheGly 120
DB 778 TTRACCGATGAGATCGAAGAGACATTCGAACTGAGTTTCCAAACCGCGCAATTCGGT 837
QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
DB 838 TTCAGCGACGACCTTAATCAGCCGAGACAGATTCCTCAAGCTGCTTCCGAACTCTG 897
QY 141 SerAsnProArgSerArgArgIleuThrArgGluGlyLeuLeuAspAspGluGluAlaThr 160
DB 898 TCTTAATCCCGCTCTGAAGAGATCAATGAAGGCTTCTTGATGATGAAGAAGCTTACA 957
QY 161 ValIleThrAspValProIleAsp----- 168
DB 958 GTCATCACTGATGTTCTCTGGATGAAGTAATTCGATTTCCGAAATTAAGTCTTCTTC 1017
QY 169 -----LysValProGlyVal 173
DB 1018 GTTTTAATTCATGAATGATGAATGAAGAAGAACTTTATCTAGTGAAGGTTCTCGGCG 1077
QY 173 AleuCysValLeuGlnGluGlyGlyGluThrGluIleValIleAspArgValGlyGluAlaLe 193
DB 1078 TCTCTGTATTCGCAAGAGAGTGATGAGACTGAGATAGTCTTCGGGTGGTGAAGCTCT 1137
QY 193 uleuLysGluArgLeuProLysSerPheLysGlnAspValValIleValIleMetAlaLeuAl 213
DB 1138 GCTTAAGAGAGAGGTCTCTTAAGTCTTTTAAGCAAGATGTGTTTATGTTATGGCGCTTC 1197
QY 213 aPheLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGly 233
DB 1198 GTTTCGATGTCGAGGAGATGCTATGCAATTCGATTCACCGATTTATTACTGCTTA 1257
QY 233 xGluPheValGluGluAlaLeuLysLeuLeuGln----- 244
DB 1258 TGAAGTTTGAAGAACTTTGAAGCTTTACAGTATGATGACTTGTGTAATTTG 1317
QY 244 ----- 244
DB 1318 ACAGCGGTGGCTTTAAAGAACTTTGATTTGATCTTTGTTATTCAGTCTTGTA 1377
QY 245 -GluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuG 264
DB 1378 GGAAGAAGAGCAAGTACCTTGCAACGATTTACGTGCAAAATTAAGAGAGCTTTGGA 1437
QY 264 uGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyrAl 284
DB 1438 AGAGATCACTCCGCGTTATGCTTGAGACTACTTGACCGCTTGAGATGATTAACG 1497
QY 284 aAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuThrSerValGlyGly 304
DB 1498 TCGGAAGAAGCTAAATGTTTAAGCGGTGCGGAATATTTGTGCTGTGGAGAGAG 1557
QY 304 yGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLe 324
DB 1558 TGAAGATCACTGCTTGTGGGGTTTGACCCGTGAAGATTATGAATGAGCGCTTTT 1617
QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
DB 1618 ATGAATGACAGCTGCTGAGAGAGGT-ATACAGTTAGATACCTTTTAAATTTCTTTAGC 1676
QY 333 -----AspLeuPheVal 336
DB 1677 ATGATTAACCTTAGGTTCTCATTTTAAATGATGTTGTGTAGCTGATCTTTTGTG 1736
QY 336 lAlaThrProSerAsnIleProAlaGluSerPheGluValIleGluValAlaLeuAlaLe 356
DB 1737 AGCTACCCCAAGCAATATTCACAGAGTCAITTTGAAGTTTACAAAGTTGCACTTCTCT 1796
QY 356 uValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPh 376

Db 1797 TGTGGCTCAAGCTTTTATTGTTGTAAGAACCAACCTTTTACAGGATGCTGATAAGCAATT 1856
 Qy eGInGInLeuGInGInLaIySValMeAlaMerGInLeProAlaMeLeuTyraAPh 396
 Db 1857 CCAAGCACTTCAGCAGGCTTAAGGTAAGCTATGAGATTCCTCGAGTGTATGATGATAC 1916
 Qy rArGAsnAsnTrpGInuIleAepPheGlyLeuGInuArgGlyLeuCySaIaLeuLeuIleG 416
 Db 1917 ACAGGAATATTGGAGATAGACTTCGCTATGAAAGGGAGACTCTGTCACTGCTTATAG 1976
 Qy yLysValaAspGInuCySaArgMeTrpLeuGlyLeuAspSerGInuAspSerGInuTyraGAs 436
 Db 1977 CAAGTGTATGATCAATGCCGTATGCTGCTTACAGACAGTGAAGATTCACAATATAGAA 2036
 Qy nProAlaIleValaGInuPheValaLeuGInuAsnSerAsnArgAspAsnAspAspLeuPr 456
 Db 2037 TCCAGCTATTGTGAGATTTGTTTGGAGAAATTCAAATCGATATACAAATGATGATCTCC 2096
 Qy oGlyLeuCySaLeuLeuGInuTrpLeuAlaGlyValaValaPheProArgPheArgAs 476
 Db 2097 TGGACTATGCAAAATGTTTGGAAACCTGCTGGCAGGGGTTCTTCTTCTAGGTTCAAGA 2156
 Qy pThrLyAspLyAspPheLyLeuGInuAspTyrrTyraAspProMetValLeuSerTy 496
 Db 2157 CACCAAAAGATATAAAATTTAACTCGGGACTACTATGATGATGATCTATGCTTGAAGTTA 2216
 Qy rLeuGInuArgValaGInuValaGInuGlySerProLeuAlaAlaAlaAlaAlaAlaAlaAla 516
 Db 496 rLeuGInuArgValaGInuValaGInuGlySerProLeuAlaAlaAlaAlaAlaAlaAlaAla 516
 Qy 2217 CTGGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2276
 Qy gIleGlyAlaGlyuHisValySAsaSerAlaMetGInaLeuGInuLySValaPheProse 536
 Db 2277 GATTGGAGCCGAGATGTGAAGCTAGTGTCTATCAGACACTGACGAAGATTTTCTTCTTC 2336
 Qy rArgTyrrThraAspArgAsnSerAlaGInuProLyAspValaGInuTrpValaPheSerVa 556
 Db 2337 CCCTATACAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2396
 Qy lAspProValaGlyAsnAsnValaGlyArgAspGlyGInuProGlyValaPheIleAlaGlyAla 576
 Db 2397 AGATCCGTGTGTGAAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2456
 Qy aValArgProSerGInuAsnPheGInuTrpAsnAspTyrrAlaIleArgAlaGlyValaSerG 596
 Db 2457 TGTAAAGCCTCTGCAAACTTTGAAACTATGATTAATGATTAATGATTAATGATTAATGATTAATG 2516
 Qy uSerSerValaAspGInuTrpThraValaGInuMetSerValaAlaAspMetLeuLySValaLe 616
 Db 2517 GAGTAAAGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2576
 Qy rValyLySValaLeuAlaAlaGlyValaAlaIleGlyLeuIleSerLeuPheSerGInuLyS 636
 Db 2577 TGTGAAGATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2636
 Qy rPheLeuLySAsnSerSerPheGInuArgLyAspMetValaSerSerMetGInuSerAs 656
 Db 2637 TTTTCTTAAAGAGCTCACTCTTTTCAAGCAGAGATATGATTTCTTCTTATGAGATCTGA 2696
 Qy pValaIaThrIle----- 660
 Db 2697 TGTGCTACCATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2756
 Qy 660 ----- 660
 Db 2757 ATGCTGTTTGTGAGCTAAGAACTATGTTCCACTTAATATACATGTCCTCAAAAGTTGTAC 2816
 Qy 660 ----- 660
 Db 2817 CAAGATTAACAAGTGTGAGTAATTTCACTAATTAATGCTGCTTGAATTTTGTATCA 2876
 Qy 660 ----- 660
 Db 2877 ACTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATAAGTGAAGTTA 2936

Qy 660 ----- 660
 Db 2937 GAGATTCCTTATGATGATGCTTTTGTCCAACTTTTCTTCTTATTTTCTTCTTATTTTCTTCTTAT 2996
 Qy 661 --GlySerValaArgAlaAspAspSerGInuAlaLeuProArgMetAspAlaArgThrAlaG 680
 Db 2997 TAGGTCAGTCAGAGCTGAGATGATTAACAGACACTTCCCAAAATGATGATGATGATGATGATG 3056
 Qy 680 IuSnIleValSerLySValaGInuLySValaGInuLySValaGInuLySValaGInuLySValaG 700
 Db 3057 AGATATATGATTCACATGTCAGAGATTAAGTCTGCTGCTTGGGCTGATCAACCGCA 3116
 Qy 700 IeGInuMetLeuProGlu----- 705
 Db 3117 TAGAATGTTTACAGAGCTGAGAGATTAATTTCAATTCATCAATTCATTCATTCATTCATTCATTC 3176
 Qy 706 -----ValLeuAspGly 709
 Db 3177 TTGACATGATTAATGATGATGCTTGTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 3236
 Qy 710 ArgMetLeuLySValaGInuTrpThraAspArgAlaAlaGInuTrpAlaGInuLeuGlyLeuValaTy 729
 Db 3237 CGATGCTGAAGATTTGAGCTGACAGACAGCTGAACCTCGCAGCTTGGGCTTGAAT 3296
 Qy 730 AspTyrrThraLeuLeuLySValaAspSerValaThrValaSerAlaAspGlyThrArg 749
 Db 3297 GATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 3356
 Qy 750 AlaLeuValaGInuAlaThrLeuGInuGInuSerAlaCyLeuSerAspLeuValaHisProGlu 769
 Db 3357 GCTCTGTGTAACCACTCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3416
 Qy 770 AsnAsnAlaThrAspValaArgThrTyrrThraArgTyrrGInuValaPheTrpSerLyS 789
 Db 3417 AACAAATGCTACTGATGTCAGAACCTTACACAAAGATATGATTTCTGCTCAAGTCA 3476
 Qy 790 GlyTrpLySValaThrGInuGlySerValaLeuAlaSer 801
 Db 3477 GGGTGAATAATCACTGAAGGCTGCTGTTTGTGATCA 3512

RESULT 7
 AP008208 011/c
 WPCOMMENT
 Sequence split into 360 fragments LOCUS AP008208 Accession AP008208
 Fragment Name Begin End
 AP008208_000 1 110000
 AP008208_001 100001 210000
 AP008208_002 200001 310000
 AP008208_003 300001 410000
 AP008208_004 400001 510000
 AP008208_005 500001 610000
 AP008208_006 600001 710000
 AP008208_007 700001 810000
 AP008208_008 800001 910000
 AP008208_009 900001 1010000
 AP008208_010 1000001 1110000
 AP008208_011 1100001 1210000
 AP008208_012 1200001 1310000
 AP008208_013 1300001 1410000
 AP008208_014 1400001 1510000
 AP008208_015 1500001 1610000
 AP008208_016 1600001 1710000
 AP008208_017 1700001 1810000
 AP008208_018 1800001 1910000
 AP008208_019 1900001 2010000
 AP008208_020 2000001 2110000
 AP008208_021 2100001 2210000
 AP008208_022 2200001 2310000
 AP008208_023 2300001 2410000
 AP008208_024 2400001 2510000
 AP008208_025 2500001 2610000
 AP008208_026 2600001 2710000

AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	10110000
AP008208_101	10100001	10210000
AP008208_102	10200001	10310000
AP008208_103	10300001	10410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000
AP008208_132	13200001	13310000
AP008208_133	13300001	13410000
AP008208_134	13400001	13510000
AP008208_135	13500001	13610000
AP008208_136	13600001	13710000
AP008208_137	13700001	13810000
AP008208_138	13800001	13910000
AP008208_139	13900001	14010000
AP008208_140	14000001	14110000
AP008208_141	14100001	14210000
AP008208_142	14200001	14310000
AP008208_143	14300001	14410000
AP008208_144	14400001	14510000
AP008208_145	14500001	14610000
AP008208_146	14600001	14710000
AP008208_147	14700001	14810000
AP008208_148	14800001	14910000
AP008208_149	14900001	15010000
AP008208_150	15000001	15110000
AP008208_151	15100001	15210000
AP008208_152	15200001	15310000
AP008208_153	15300001	15410000
AP008208_154	15400001	15510000
AP008208_155	15500001	15610000
AP008208_156	15600001	15710000
AP008208_157	15700001	15810000
AP008208_158	15800001	15910000
AP008208_159	15900001	16010000
AP008208_160	16000001	16110000
AP008208_161	16100001	16210000
AP008208_162	16200001	16310000
AP008208_163	16300001	16410000
AP008208_164	16400001	16510000
AP008208_165	16500001	16610000
AP008208_166	16600001	16710000
AP008208_167	16700001	16810000
AP008208_168	16800001	16910000
AP008208_169	16900001	17010000
AP008208_170	17000001	17110000
AP008208_171	17100001	17210000
AP008208_172	17200001	17310000

QY 333 -----AspLeuPheValA 337
 DB 80944 CATCTTGAACAACTGTAACCTCCCTTCTGTACACCTCCAGATGATTTCTTTTCAA 80885
 QY 337 laThrProSerAenIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeu 357
 DB 80884 AAACACGGAATAGCATCTCTCGTAATGTTGAAATTTTACAAATGTGACACTTGACATG 80825
 QY 357 alaAlaGlnAlaPheIleGlyLysLysProHISleuLeuGlnAAspLysGlnPheG 377
 DB 80824 TCCCTCAAGCAATTAATAGTAAGCCCAATTCATCATGATGACGATGATCTTTTGG 80765
 QY 377 lngInleuGlnGlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrA 397
 DB 80764 AACCACTCCAGAAAGTTACACATGAGTTCT-----CATTAAGCTTAATGATTAAT- 80718
 QY 397 rgaAenIleTyrGluIleAspPheGlyLeuGluIleGlyLeuCyAlaLeuLeuIleGlyL 417
 DB 80717 -----GAGATGACCTTGCTCATTTGGAAGGCGCATTTCTGCTCATTTGCTGCGAG 80669
 QY 417 ysaValaAspGluCyAlaMetTyrLeuGlyLeuAspSerGluAspSerGlnTyrArgAsp 437
 DB 80668 ATGTTAGCAAGTCAGATGTCCTTGAATGATTAATGATGATCTTCAACATACAGAGACC 80609
 QY 437 roAlaIleValaGluPheValleuGluAenSerAen---ArgAspAspAspAspLeuP 456
 DB 80608 CCAAAATTTAGAGTTATTTGTGACCAACTTACATCATGATGAGAGAAATGATCTTTTTC 80549
 QY 456 roGlyLeuCyAlaLeuLeuGluIleTyrPheValaGlyValaPheProArgPheArgA 476
 DB 80548 CAGGGCTGTGCAAGCTTTTGAAGCTTGGCTTATCTTGAAGCTTTTTCCTAGAGCAGAG 80489
 QY 476 spThrIlyAspLysLysPheLysLeuGlyAspTyrTyrAspAspPheMetValleuSert 496
 DB 80488 ATACTCGGGCATGACGTTGACACTTGAAGATTACTACGATGATCCAGAAATTTTAAAGCT 80429
 QY 496 yrlleuGluIlyValaGluValaGlnIleGlnIleSerProleuAlaAlaAlaIleThrMetAlaA 516
 DB 80428 ACCTAGAAAGATGAGAGGCTGCTGCTCTTCTCATTTGGCTGCTGCTGCTGCTATTCGA 80369
 QY 516 rglIleGlyAlaGlu-----HisValaLysAlaSerAlaMetGlnAlaL 530
 DB 80368 AACTGTGCTCAAGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 80309
 QY 530 eugIlyValaPhePro---SerArgTyrThrAspArgAenSerAlaGluProLysAspV 549
 DB 80308 TCAACAAAGTTTTCATGATAGACAGTTAGACAGCTCAGCCATGGAATTAATCT- 80253
 QY 549 alGlnGluIlyThrValPheSerValaAspProValaGlyAenAenValaGlyArgAspGlyGluP 569
 DB 80252 -----AAGATGGC---C 80243
 QY 569 roGlyValaPheIleAlaGluAlaValaArgProSerGluAenPheGluIlyThrAspAspTyrA 589
 DB 80242 CTGGGGGATTAATCTT-----GAAATTTTGAACAG----- 80214
 QY 589 laIleArgAlaGlyValaSerGluSerSerValaAspGluIlyThrValaGluMetSerValaA 609
 DB 80213 -----GAAATGACACCTG 80201
 QY 609 laAspMetLeuLysGluAlaSerValaLysIleLeuAlaAlaGlyValaAlaIleGlyLeu 629
 DB 80200 CTGATGATTCGAAATATCCGCTTGAAGATTATCTTCTGCGGCACTGTTGACATGCT 80141
 QY 629 leSerLeuPheSerGlnLysTyrPheLeuLysSerSerSerSerPheGlnArgLysAspM 649
 DB 80140 TGGCAGATTAATTTGGGCGCAATTAAT-----TTGCTCTGTAAAGGC 80102
 QY 649 etValaSerSerMetGluSerAsp-----ValaIleThrIleGlySerValaArgAlaA 666
 DB 80101 CCTTCTGTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80042

QY 666 spAspSerGlu----- 669
 DB 80041 CTGATGTGATGAATATGCTTTTCATTGCTAAAGTAGTCAAAACAACACTGGAATTTACTG 79982
 QY 669 ----- 669
 DB 79981 CTCACGTGTTTACTGGGTAATCTTGTGGCTGATTAACAGTAATGTTCTTATAGCT 79922
 QY 669 ----- 669
 DB 79921 TTGCATTATGTTTAACAAGCTATACATCAAAATATTTTGTGTATACATPACAGATG 79862
 QY 669 ----- 669
 DB 79861 CTACATAATCTGTAT 79802
 QY 670 -----AlaLeuProArgMetAspAlaArgThrAlaGluA 681
 DB 79801 ATCTTGCACTAGATGAAGATCCAGTACATATTTCTGAATGATGGAAGCTGGCAGAG 79742
 QY 681 snIleValaSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHISArgIleG 701
 DB 79741 ATATTTTGGCAAGTGGCAGAGATTAATCTTAAGGCTTGGGACCAAGAACTTGGGTTG 79682
 QY 701 lueMetLeuProGlu----- 705
 DB 79681 CATCATTCGAAGAGGTATGATTTTAAACTTTGACTTTTCATGCTCTCTAGCTGGCCACTG 79622
 QY 705 ----- 705
 DB 79621 CTCTTACTATACCTACGCTTCTCTATAGAGATTAATTCAGATATACTCTCTATAGGA 79562
 QY 705 ----- 705
 DB 79561 TATATTTGTCACGCAATGTAATCTTCAAGTTTGGTTACAATAATCTTTCTTTGTTGACT 79502
 QY 705 ----- 705
 DB 79501 CACACATGCTTCAATTCGAACTGAACAAATTTTCAAGAACTATTTGCCACTGCATGATA 79442
 QY 706 -----ValleuA 708
 DB 79441 TCCCGGTGATGAGCGGCTCAATGCTGTGATGATTTTCCATGTTCTTATAGTCTTGG 79382
 QY 708 spGlyArgMetLeuLysIleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuV 728
 DB 79381 ATGGCAACATGCTTAAGAGTGTGACTGACCCGACGACGAGATGAGCGTCAAGGGTGGT 79322
 QY 728 alTyrAspTyrThrIleuLysLysLeuSerValaAspSerValaIlyThrValaSerAlaAspGlyT 748
 DB 79321 TCTGGAGATATACATATCCGATGACGATGATGATGATGATGATGATGATGATGATGATG 79262
 QY 748 htrGAlaLeuValaGluAlaIleThrLeuGluGluSerAlaCyLysLeuSerAspLeuValaHISP 768
 DB 79261 GACGACGCACTGGGAGGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATG 79202
 QY 768 roGluAenAenAlaIleAspValaArgThrTyrThrThrArgTyrGluValaPheTyrSert 788
 DB 79201 CCAGAAACATATGATTAATAGACACAAATATCACTACCGGATGATGATGATGATGATGATG 79142
 QY 788 ys---SerGlyTyrLysIleThrGluGlySerValaLeuAlaSer 801
 DB 79141 AGCTAGAGAGGTGGAAGATTAACGGAAGAGCACTCTCAAGTGG 79098
 RESULT 8
 AP004885/c 150462 bp DNA linear PLN 15-SEP-2004
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 DEFINITION PAC clone: P0575F10.
 ACCESSION AP004885
 VERSION AP004885.3 GI:41053009
 KEYWORDS Oryza sativa (japonica cultivar-group)

ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryote; Vriidiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa niponbare (Ga3) genomic DNA, chromosome 2, PAC
clone: P0575F10
JOURNAL
PUBLISHED ONLY IN Database (2002)
2 (bases 1 to 150462)
AUTHORS
TITLE
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: sasaki@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Jan 21, 2004 this sequence version replaced gi:38142429.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark (http://opal.biology.gatech.edu/GeneMark/), Glimmer
(http://www.cigr.org/cdb/glimmer/gimr_form.html), RiceHMM
(http://rpg.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologues of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from T7 to SP6 of the PAC clone.
This sequence of P0575F10 clone has an overlap with OJ1020.C02
(DBJ: AP004078) clone at 5' end and with P0482F12 (DBJ: AP005311)
clone at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rpg.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1. 150462
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Niponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0575F10"
/complement(4987..8112)
/gene="P0575F10.1"
complement(4987..5520..5679..5764..5914,
6096..6247..6738..6859..7864..8112))
/gene="P0575F10.1"
/note="supported by full-length cDNA(s): AK106553"
complement(join(5003..5412..5520..5679..5764..5914,
6096..6247..6738..6859..7864..8112))
/gene="P0575F10.1"
/note="supported by full-length cDNA(s): AK073059"
complement(join(5315..5412..5520..5679..5764..5914,
6096..6247..6738..6859..7864..8112))
/gene="P0575F10.1"

/note="contains EST(s): AU070160(B61534)
contains full-length cDNA(s): AK106553, AK073059"
/codon_start=1
/product="putative eukaryotic translation initiation
factor 3, subunit 1 alpha, 35kDa"
/protein_id="BAD07941.1"
/db_xref="GI:41053010"
/translation="MEDWSDSEDFOPAPVAPSAKAEPLKMADEVEDVEDVDEWEESE
EKRPPEVETAPKPSGKAQKQASSEVYQDEALDDPLLEKROORLYEAD
FKSTTEFGCKGSEKSLDPFIRKSESDPAVEYELINKRIPREKSGYLKNWR
LSNASKGADAKDSSITAIINAEKTAEKBAAGKKQAKKQILHTEKDDPFITG
KGNFDDDEYDFM"
complement(10186..13754)
/gene="P0575F10.2"
complement(join(10186..10479..10759..10892..11127..11991,
12089..12349..12515..12712..12843..13073..13455..13754))
/gene="P0575F10.2"
/note="start and end point are not identified"
complement(join(10186..10479..10759..10892..11127..11991,
12089..12349..12515..12712..12843..13073..13455..13754))
/gene="P0575F10.2"
/note="contains EST(s): AU058423(B51136), AU183658(B51136)"
/codon_start=1
/product="plastid division protein precursor"
/protein_id="BAD07942.1"
/db_xref="GI:41053011"
/translation="MGEFHNLRLPNSAPPAFAFSLPRPRPRRRPPPSACRAAR
MARLRPADPHLPTAAPSDPSAPAPAAAPSAFPPPLPPDAERSLPQVDYTKL
GAPHFAGDGRFEARLKPQVGTDLALGRRMQLIADTLNKSRTYDRA
LSNRREGLTMDIAMDKEGALAVATGSQLDRPKKFDVAVLAMAAYDLR
DMAASPPVIGCCLELRLALQLEBAKALPDLSQIDETLEETPCVLELIL
PITENHKKKQOEGOGARNTLWVGSGIATVGGFGRBAFMNBAFLRMSIEGMDF
MDLRGSPPEPPIYVALAHVAADILISRSPPITMADLIPEOLKRNIGSHVAYNE
KTLALERAFCSLLVGVSKRCMVLGIDNESSPTPRDLTFITNNSISENDLPL
CKLETWLFEVPRSRDTRGMQFRLDYDDDEVLSLTEREGGASHLAAALAIK
LGQATLALCTVSNIAIQANKFPLEIQDRAMENTKGGPGVLENFQENABAD
SRNAALKTISGALPALAVITGAKYLPKRLPAISEHGSVAVANSVDTPDALD
DPAHIFPMADKLAEDIVRKQSIKSAVLGSEHSAVLAQEVLDGNMMLKVMIDRAAIR
HGWFWYTLSDATYDITSDTSLDGRRAVETIDEAQLDVTETPRNDSVDTKTTRY
EMAFSLGKGWKTIGAVLKS"
complement(16401..17341)
/gene="P0575F10.3"
complement(join(16401..16739..16841..17341))
/gene="P0575F10.3"
/note="start and end point are not identified"
complement(join(16401..16739..16841..17341))
/gene="P0575F10.3"
/note="similar to Arabidopsis thaliana chromosome 3,
At3g61870"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07943.1"
/db_xref="GI:41053012"
/translation="NMAAPLAVAHAIITCSASNKNSPPSARQQOCTTTATGSPAA
LPSELRTTAAATTAALALAPALAAAGGFTLVBGSVALHPLVWGAFATYLVG
YLGWKRVRVTTODETINELKOLKPAALAAATPAVAAAGSSSSPPSPAPSVPI
KIDELTBERRKLIKSGFRDRHFNAGSILILGCTBSYGCLNTMFRGKLPFGPHLPA
GAATVLMMAAALVPMQKGNETARSILHAIINVLPIWQIPGLSLVGVKFEFT
TWP"
complement(18499..18942)
/gene="P0575F10.4"
complement(<18499..>18942)
/gene="P0575F10.4"
/note="start and end point are not identified"
complement(18499..18942)
/gene="P0575F10.4"
/note="contains EST(s): AU184851(R10374), AU184850(R10374)"
/codon_start=1
/product="putative EF-hand Ca2+-binding protein CCD1"
/protein_id="BAD07944.1"
/db_xref="GI:41053013"
/translation="NMHQQIRSTTAABQOESAGGGGEBYEDLMPVAGRLGAGCL
LSELRAQFRLLADPARCAITAEISLRSAASVILGIGGGGEMTEBAAMVREBDQGD

gene
21937..23552
/gene="P0575F10.5"
join(<21937..22225,22326..22648,22736..23049,
23393..23552)
/gene="P0575F10.5"
/note="start and end point are not identified"
join(21937..22225,22326..22648,22736..23049,23393..23552)
/gene="P0575F10.5"
/note="similar to Arabidopsis thaliana chromosome 5,
At5g24330"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07945.1"
/db_xref="GI:41053014"
/translation="MGPRATPLRRRTARRAPATRAEGSGDGDNDVRCAGCGGSA
AELLDCDCGRGHIIFCLRPILPRVPGRDFPCSPASPSHKSMAKPKPPVLO
TKIVDFPKIQRGPAALAAAESEEGKRRKVGIRLVSKRRKLPERSPDDPAR
LRMASLATATATGAVFSNELTYVPGMARANRALESQGVLPKEDEVTANLCK
RMAARGMPPLVVDVEGTVADRFIDLTITVYGVDDVLTREHDDGSMWT
LSAATPSRSLVCPDKRSNIARFINGINHTPDGRKKQMLKCVRPDVGECVLLVA
NRDISGRLTYDNGSEHEYPTHHPV"
complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/gene="P0575F10.6-1"
complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28147))
/gene="P0575F10.6-1"
/note="supported by full-length cDNA(s): AK121372"
complement(join(24491..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28147))
/gene="P0575F10.6-1"
/note="supported by full-length cDNA(s): AK062255"
complement(24563..28147)
/gene="P0575F10.6-2"
complement(24563..28147)
/gene="P0575F10.6-2"
/note="contains full-length cDNA(s): AK065148
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
complement(join(24787..24817,24913..25266,25398..25633,
25720..25832,27054..27185))
/gene="P0575F10.6-1"
/note="contains EST(s): AU031017(E60571), AU093466(E60571)
contains full-length cDNA(s): AK121372, AK062255"
/codon_start=1
/product="putative arginine/serine-rich splicing factor"

Alignment Scores:
Pred. No.: 1,79e-86 Length: 150462
Score: 1540.50 Matches: 398
Percent Similarity: 41.7% Conservative: 118
Best Local Similarity: 32.2% Mismatches: 215
Query Match: 37.9% Indels: 505
DB: 15 Gaps: 20

US-10-600-070b-2 (1-801) x AP004885 (1-150462)

QY 12 SerProhGlnLeuCyvArgLeuProAlaThrThryLeuAlaGArgSerHisAan 31
:::|||||
DB 13715 GCGCAATTCCTTCCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCT 13656
:::|||||
QY 32 ThSerThrThrIleCyS---SerAlaSerIleThrAlaAspArgLeuLeuSerAspHe 50
:::|||||
DB 13655 CACCCCTCCGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCTTCGCGCACTTC 13596
:::|||||
QY 51 AsnPhe-----ThSerAspSerSerSerSerSerPheAlaThrAlaThr 65
:::|||||
DB 13595 CACTCTCTCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 13536
:::|||||
QY 66 ThrThraAlaThrIleuValSerIleuProSerIleAspArgProGluArgHisValPro 85
:::|||||
DB 13535 GCGCGCTCCGCGCTCCCTTCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCC 13476
:::|||||

QY 86 IleProIleAspPheTy----- 91
:::|||||
DB 13475 CTCGAAGTGATTTCTTCAAGAGGAGGAGCCCGGAAATCTGTCGGTTTGTTCATCATTT 13416
:::|||||
QY 91 ----- 91
DB 13415 GGAAGCTCGAATATCATCCCGTAGATMAAAGGAGATGATTTCTTCTTGTGTAG 13356

QY 91 ----- 91
DB 13355 GTGATGATACATTTGATCTGCTTAATGCTTGAAGCTTATTCACCTGCAGCTTTTGC 13296

QY 91 ----- 91
DB 13295 TAGGTTACTGATCCCTCAGTTCATGAATTCAGATTATTGCTATTCCGTAGTGCATC 13236

QY 91 ----- 91
DB 13235 TCAGTAGTTATTACTGATCAATGCGAATCCATCGCATTTGTGATCCACTGCATAC 13176

QY 91 ----- 91
DB 13175 TATGATACGAATTATGCTTATCATGTGTGTGGGCTAGAGATCATGCTTTT 13116

QY 92 ----- 92
DB 13115 TGACATATTCGAAATGTTTGTGTAATCTGCTTGCAGGTTCTAGGGGAGAGACCA 13056

QY 99 HisPheLeuThrAspGlyIleArgAlaIlePheGluAlaArgValSerIleProProGln 118
|||||
DB 13055 CATTTCTTGGCCATGATCATGAGAGGGGCTTCGAGCAACGATCCAGCCGCGCAG 12996
|||||
QY 119 PheGlyPheSerAspAspAlaLeuIleSerArgArgGluIleLeuGlnAlaIleCysGlu 138
:::|||||
DB 12995 TATGCTACAGACCGAGATGCTCTTGTGCTCGCAATGCTGAGATGCCCATGAC 12936
|||||
QY 139 ThrLeuSerAsnProArgSerArgArgArgGluIleArgGluIleLeuLeuAspArgGlu 158
|||||
DB 12935 ACTCTGATGACAGCAAGACTCCGCACTGATGATGATGATGATGATGATGATGATGAT 12876
|||||
QY 159 AlaThrValIleThrAspValProTrpAspIleY----- 169
:::|||||
DB 12875 GAAGCTCTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12816
:::|||||
QY 169 ----- 169
DB 12815 GTGTTTGATCTTGTCAAAATGACACTAGTGTATATCTCTCTGCTTATATATATG 12756

QY 170 ----- 170
DB 12755 GTGCTTTGAGACAGAGTTCCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 12696

QY 184 uIleValLeuArgValIleGluAlaLeuLeuIleGluAlaLeuLeuProIleSerIlePheIle 204
|||||
DB 12695 TGCTGTGCTTGTATCTGAGAACAGTGTCTTGTGATCGGCCCAAGGCTTCAAGCA 12636
|||||
QY 204 HisPheValIleValIleMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLe 224
|||||
DB 12635 GGAAGTGTGTCTGACAGATGCTGTGCTTATGTGATCTATCAAGGAGATCTATGGCAGC 12576
|||||
QY 224 uAspProAspPheIleThrGlyIleArgIleValIleGluAlaLeuLeuLeuLeuIle 244
|||||
DB 12575 AACCCCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12516
|||||
QY 244 n----- 244
DB 12515 GTGAACTGTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 12456

QY 244 ----- 244
DB 12455 AGATTAAATACCTGTGAAAGATGTTGTTGGACACATGGGCTATATTTTAAATAGCA 12396

Qy 245 -----gluInglYAlase 249
 Db 12395 ATGCGATTTCTTGAAACATATTAAATAGAACTTCACTGATGAGAAAGTGAAGCAAG 12336
 Qy 249 rSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGluGluIleThrProAr 269
 Db 12335 CAATCTGCGACCTGATCTGCTTTCACAGATTGATGAACCTCTCGAGAGATTACACCTCG 12276
 Qy 269 gTyValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyValAlaAlaValAspLeuAs 289
 Db 12275 CTGTGATTTGAGCTTCTCTCCCTTATTTGACACAGCATCATGAAGACCGCCAGAGA 12216
 Qy 289 ngIleuSerGlyValArgAsnIleLeuThrPheValGlyGlyValAlaSerAlaLe 309
 Db 12215 AGGCTTCAAGCTGCGGAAACATTTTGTGAGCTTGGCAGAGAGAGATTGCTTACCCT 12156
 Qy 309 uValGlyLeuThrArgGluTyPheMetAsnGluAlaPheLeuArgMetThrAlaAl 329
 Db 12155 TGGAGGAGATTTTCTGTGAGCCCTCATGAACGAGCTTTTGTGAGATGACATCAAT 12096
 Qy 329 agIuGlnVal----- 332
 Db 12095 TGAACAGTTCTGATCATGTTGTTTATGAGACAGCTGCTTATTTATGACAGACA 12036
 Qy 333 -----AspLeuPheVala 337
 Db 12035 CATCTTGAACAACTGTTAACTCCTTCTGTTACACTCCAGATGATTTCTTTTCAA 11976
 Qy 337 laThrProSerAsnIleProAlaGluSerPheGluValTyArgIuValAlaLeuAlaLeu 357
 Db 11975 AAACACCGAATAGCATCTCTCGAATGTTGAAATTTTCAATGATGACATTCGACATG 11916
 Qy 357 alaIleGlnAlaPheIleGlyValLeuProHISleuLeuGlnAspAlaAspIleGlnPhe 377
 Db 11915 TCCTCTAAGCAATATATGATGAAGGCCAACATTCATGATGCGGATGATCTTTTGG 11856
 Qy 377 lngInleuGlnGlnAlaValMetAlaMetGluIleProAlaMetLeuTyAspThr 397
 Db 11855 AACAACTCCGAAAGTTCAACATGAGTTCT-----CATTAAGCTTATGATTAAT- 11809
 Qy 397 rgaAsnAspTrpGluIleAspPheGlyLeuGluArgGlyLeuCyalaLeuLeuIleGly 417
 Db 11808 -----GAGATGACCTTGCACTTGGAAGGCAATCTGCTCATCTGCTAGTCGAG 11760
 Qy 417 ysValAspGluCyalaMetTrpLeuGlyLeuAspSerGluAspSerGlnTyArgAsp 437
 Db 11759 ATGTTCAGCAAGTGAATGCGCTTGAAATGATATGATGATCTTCCATACAGAGACC 11700
 Qy 437 roAlaIleValGluPheValLeuGluAsnSerAsn-----ArgAspAspAspAspLeu 456
 Db 11699 CCAAAATCTTAGAGTTATTGTGACCAACTTACATCACTGACAGTGAAGAGAAATGATCTTTC 11640
 Qy 456 roGlyLeuCyalaLeuGluGluThrTrpLeuAlaGlyValAlaPheProArgPheArg 476
 Db 11639 CAGGGCTGTCAAGCTTTTGGAGACTTGCTTACTTGAAGTTTTCCTAGAGAGCAGAG 11580
 Qy 476 spThrIleAspIleValysPheLeuGluGlyAspTyTyArgAspAspPheMetValLeuSer 496
 Db 11579 ATACTCGGGCAGTCAGCTTCACTTGAGATTACTAGATGATCCAGAGTTTAAAGCT 11520
 Qy 496 yTrleuGluArgValGluValAlaGlnIleSerProleuAlaAlaAlaAlaAlaAla 516
 Db 11519 ACCCTAGAAAGATGAGAGGCTGCTGCTTCTCATTTGGCTGCTGCTGCTGCTATTCAA 11460
 Qy 516 rgIleGlyAlaGlu-----HisValIleValAlaSerAlaMetGlnAla 530
 Db 11459 AACTGTGCTCAACGATCAAGCTGCACTGTGATCTGGAATCAAAAGCTATTCACAGCT 11400
 Qy 530 euGlnIleValPhePro-----SerArgTyrThrAspArgAsnSerAlaGluProIleAsp 549
 Db 11399 TCAACAGAGTTTTCATTGATAGAACATTTAGACAGCTCAGCATGAAATAACT- 11344
 Qy 549 alGlnIleuThrValPheSerValAspProValGlyAsnAsnValGlyArgAspGlyLeu 569

Db 11343 ----- 11334
 Qy 569 roGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAspAspTyra 589
 Db 11333 CTGGGGATATCTT-----GAAAATTTTGACAG- 11305
 Qy 589 laIleArgAlaGlyValSerGluSerValAspGluThrThrValGluMetSerVala 609
 Db 11304 -----GAAAATGCACCTG 11292
 Qy 609 laAspMetLeuTyGluAlaSerValIleValAlaAlaGlyValAlaAlaIleGlyLeu 629
 Db 11291 CTGATGATTTGAGAAATGCCCTTGAAATTTATCTGCTGGGCACTGTTGCACTGT 11232
 Qy 629 leSerIleuPheSerGlnIleTyThrPheLeuTySerSerSerSerPheGlnArgIleAsp 649
 Db 11231 TGGCAGTAAATTTGGGCGCCAAATAT- 11193
 Qy 649 etValSerSerMetGluSerAsp-----ValAlaThrIleGlySerValArgAla 666
 Db 11192 CCTTCTGCTATTTAGAGTGAAGCATGATCTGTGCGATTTGCTAATGATGACTCTA 11133
 Qy 666 spAspSerGlu----- 669
 Db 11132 CTGATGTGAGTAAATGCTTTTCAATTGCTAAAGTAGTCAACACACTGGAATTTACTG 11073
 Qy 669 ----- 669
 Db 11072 CTCACGTTTACTGGGTATCTTGTGGCTGATTAACATGATGTTCTGATTTATGACT 11013
 Qy 669 ----- 669
 Db 11012 TTGCATTAATGTTTAACAAAGCTAATACATCAAAATTTTGTCTCATACATACAAATG 10953
 Qy 669 ----- 669
 Db 10952 CTACTAATATTCGTATATGATATATCATAGAGTTAACTTTGCATGATGATGATTTGAAG 10893
 Qy 670 -----AlaLeuProArgMetAspAlaArgThrAlaGlu 681
 Db 10892 ATCTGCACTAGATGAAGATCCAGTACATATTTCTTGAATGATGCGAACCTGGCAGAG 10833
 Qy 681 snIleValSerIleTrpGlnIleValIleSerIleuAlaPheGlyProAspPheArgIleG 701
 Db 10832 ATATTTGCGCAAGTGCAGATCATTAATCTTAAGCTTGGACCAAGCAATTCGGTTG 10773
 Qy 701 lueMetLeuProGlu----- 705
 Db 10772 CATCATTTGCAAGAGGATGTTTAAACTTTGACTTTTCACTGCTTCTACTGCGCACCTG 10713
 Qy 705 ----- 705
 Db 10712 CTCTTACCTATACCTAGCTTCTCTATAGAGATTAATTCAGATATACTCTCTATAGGA 10653
 Qy 705 ----- 705
 Db 10652 TATATTGTCAGCAATGATCTTTTCAAGTTTGTTCACAAATATCTTTCTGTTGACT 10593
 Qy 705 ----- 705
 Db 10592 CACACATGCTTCAATTGCACTGAACAAATTTTCAAGAACTTATGCCACTGCATGATATA 10533
 Qy 706 -----ValIleu 708
 Db 10532 TCCCGGTAGATAGCGGCTCAATATGCTGATGATTTTCCATGTTCTTATAGGTTCTTG 10473
 Qy 708 spGlyArgMetLeuTyIleTrpThrAspArgAlaAlaGluThrAlaGlnIleGlyLeu 728
 Db 10472 ATGGCAACATGCTAAAGTGTGACATGACCGACAGCGGAGATTGAGCGTCATGGGGTGT 10413
 Qy 728 alTyAspTyThrIleuLeuTyLeuSerValAspSerValThrValSerAlaAspGlyT 748

Db 10412 TCTGGAGATATACATTCGATGTGACGATGTATGACATCACTATCTCCCTAGATGCTC 10353
Qy 748 hrarqalaleuValgluAlaThrleugluGluSerAlaCySeuSerAspLeuValHis 768
Db 10352 GACGAGGAGCTGCTGAGGCTTCACTTATGATGAGGAGCCAACTTACGATCTTACTGAC 10293
Qy 768 rogluenaenAlaThrAspValArgThrThyThrThrArgTyrgluValPheTrpSerL 768
Db 10292 CCAGAAACATGATTCATGATGACACAAATAACTACCTACCCGGTATGAGATGAGCCTTCCTCA 10233
Qy 788 yS---SerGlyTyrPlyValIleThrGluGlySerValIleuAlaSer 801
Db 10232 AGCTAGAGGGGTGAGATTAACGAGAGGACGATCTCTCAACTCG 10189
RESULT 9
AC158210
LOCUS Medicago truncatula clone mch2-155019, WORKING DRAFT SEQUENCE, 4
DEFINITION Medicago truncatula clone mch2-155019, WORKING DRAFT SEQUENCE, 4
AC158210 133779 bp DNA linear HTG 22-JUL-2005
ordered pieces.
AC158210
AC158210.11 GI:71061528
HTG: HTG PHASE2; HTG DRAFT
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Medicago truncatula BAC Clone mch2-155019
JOURNAL Unpublished
2 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Jul 22, 2005 this sequence version replaced gi:68342147.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 43719: contig of 43719 bp in length
* 43720 43819: gap of unknown length
* 43820 83929: contig of 40110 bp in length
* 83930 84029: gap of unknown length
* 84030 106272: contig of 22243 bp in length
* 106273 106372: gap of unknown length
* 106373 133779: contig of 27407 bp in length.
Location/Qualifiers
1..133779
/organism="Medicago truncatula"
/mol_type="genomic DNA"

/db_xref="taxon:3880"
/clone="mch2-155019"
/clone_lib="Medicago truncatula BAC library H2"
43720..43819
/estimated_length=unknown
83930..84029
/estimated_length=unknown
106273..106372
/estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.: 1,82e-67 Length: 133779
Score: 1241.00 Matches: 319
Percent Similarity: 33.9% Conservative: 63
Best Local Similarity: 28.3% Mismatches: 87
Query Match: 30.5% Indels: 659
DB: 14 Gaps: 9
US-10-600-070B-2 (1-801) x AC158210 (1-133779)
Qy 7 ValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuProAlaThrThyLeu 26
Db 130423 ATGGGCTCAGCTCTTCCACTTCCG-----CGTTCACAGCTTCTCCAAAACCT 130476
Qy 27 ArgArgSerHisAnThrSerThrThrIleCySerAlaSerIySTrpAlaAspArgLeu 46
Db 130477 AACCGTCCAT-----TCCTCGCGGTCTCCGCCACGATGAATGCGGAGCATCTC 130530
Qy 47 LeuSerAspPheAsnPhethrSerAspSerSerSerSerSerhelathrAlaThrThr 66
Db 130531 ATTTCCGATTTCCAAATTTCTCGGCGACACCTCTCTCTCTCCACGACACCTCCGCC 130590
Qy 67 ThrAlaThrLeuVal---SerLeuProProSerIleAspArgProGluArgHisValPro 85
Db 130591 ACAGTCACTCTCACTCTCTTTCCTCTCCGATA-----GAACCCACGTGTCTA 130641
Qy 86 IleProIleAspPheThyrgluValIleuGlyAlaGlnThrHisPheLeuThrAspGlyIle 105
Db 130642 CTCCTCTCGACCTGTACAAATTCCTCGGCGCGAAGCATTTCTCGGTATGATAT 130701
Qy 106 ArgArgAlaPheGluAlaArgValSerIyProProGlnPheGlyPheSerAspAspAla 125
Db 130702 CCGAGAGCTTATGACGAATTCGAAAGCCCTTCAGATGTAAGAAGCT 130761
Qy 126 LeuIleSerArgArgGlnIleLeuGlnAlaIaCyGluThrLeuSerAspProAspSer 145
Db 130762 TTGATTAGTCGTCTGACGATTCCTCAAGCTGTGTGAAACCTTACCTGATCTGCTTCT 130821
Qy 146 ArgArgGluTyrAsnGluGlyLeuLeuAsp-----AspGluGluAlaThrVal 161
Db 130822 AGAAGAGATATATATCAAGCCCTCGACGACGATGAAGACGAAGATAGAGGATCTTCAT 130881
Qy 162 IleThrAspValProTrpAsp-Lys----- 169
Db 130882 CTCACGAAATCCCTTTCGACAAAGTAAAGCTTCAATTCATTCATTCATTCACATTT 130941
Qy 169 ----- 169
Db 130942 GAGTTTGTCTTCATTAATAAAGATTAAATGTTGTTAATTCGAAATGATATATGAGTTTCA 131001
Qy 169 ----- 169
Db 131002 TTTTACTAAAACACTTTTATTAGTGAAGTACTAGTATGAGATTAAGATTAAGATAAATAAT 131061
Qy 169 ----- 169
Db 131062 AGTTTAAGACATGTTTTCGAAAAACAGTGAATAGTTGATTAAGACACTTATAGTGATGA 131121
Qy 169 ----- 169
Db 131122 GCTTATGAAGATTAAGCTAGATGATTAATTTGTAGTGTGTTGTAACCTTAGCGGTTAAAC 131181

Db	133341	ACCTGAGCAATATTCACGCTGAAAGTTTGAAGCGCTATGGGGTTGACCTTGACATGGTTGG	133400
Qy	358	AGlnaApehIeGlyLybLyseProhIseLeuEngIaApeIaApeLyseGlnApeGlnG	378
Db	133401	ACAAACCTTTGAGGTAAAGACCATCTTATCCAAAGATGCTGATATATTATTCATCA	133460
Qy	378	nLeuGInGlnAlaLybAlaMetAlaMetGulIeProAlaMetLeuTyAAsp-----Th	396
Db	133461	ACTTCACAACTAAGTAAGTAACAATATATGAGAAATGCTCCCTCTGTTTATCTCCATGGA	133520
Qy	396	rArgAenAenTfpgIuIeAaSPheGlyLeuGlnAArgIyLeuCyAlaLeuLeuIeG	416
Db	133521	CATGAGAGAGAGAGAAAGTTCATTTTCATTTAGAAAGGGCTGTGTGTGCATGCTTTTGG	133580
Qy	416	LybAlaApeGlnCyAArgMeTfLeuGlyLeuApeSerGlnAAspSerGlnTyArgAs	436
Db	133581	GGAGCTTGATCATATGTCATGTCATGTTGGGTCTCGATATGACAGCTCACCTTATCGAAA	133640
Qy	436	nProAlaIeValGluPheValLeuGluAenSerAmaAArgAspAspAmaAspLeuPr	456
Db	133641	CCCACTATTAATAGACTTATATATGAGAAACCGCAAAAGGTAGAGAAAGACAGTATCTTC	133700
Qy	456	oGlyLeuCybLybLeuLeuGluTyTrpLeuAlaGlyValAlaIaPheProArgPheArgAs	476
Db	133701	TGGACTCTTAATTTGTTGGAGACATGCTTGATGAGAGCTGTGTTTCCCTTAGGTTTAGAGA	133760
Qy	476	pThrLybApeLybLysePhe	482
Db	133761	TACTAAAGACAACTT	133779
RESULT 10			
LOCUS	AC155597	184752 bp	DNA
DEFINITION	Zea mays strain B73 clone ZMMB0216X08, ** SEQUENCING IN PROGRESS		
ACCESSION	AC155597.2	GI:58082456	
VERSION	AC155597.2	GI:58082456	
KEYWORDS	HTG; HTGS; PHASE1.		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 184752)		
AUTHORS	Chan, A.P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecherback, T.R., Feldblum, T.V., Rabinowicz, P., Frazer, C.M., Schbert, K., Samiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T., Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J., and Quackenbush, J.		
TITLE	Consortium for Maize Genomics - BAC skim sequencing and assembly		
JOURNAL	Unpublished		
AUTHORS	2 (bases 1 to 184752)		
TITLE	Chan, A.P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecherback, T.R., Feldblum, T.V., Rabinowicz, P., Frazer, C.M., Schbert, K., Samiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T., Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J., and Quackenbush, J.		
JOURNAL	Direct Submission		
AUTHORS	Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850		
REFERENCE	3 (bases 1 to 184752)		
TITLE	Chan, A.P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecherback, T.R., Feldblum, T.V., Rabinowicz, P., Frazer, C.M., Schbert, K., Samiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T., Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J., and Quackenbush, J.		
JOURNAL	Direct Submission		
AUTHORS	Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850		
COMMENT	On Jan 25, 2005 this sequence version replaced gi:57863118.		
	----- Trace submission		
	Seq. name: TIGR		
	Seq. lib id: ZG30		

```

----- Project information
Web site: http://www.tigr.org/tbi/maize/
Contact: maize@tigr.org
-----
** NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown..
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*
* 1 2210: contig of 2210 bp in length
* 2211 9941: contig of 7631 bp in length
* 9942 10041: gap of unknown length
* 10042 16746: contig of 6704 bp in length
* 16746 16846: gap of unknown length
* 16846 32394: contig of 15549 bp in length
* 32395 32484: gap of unknown length
* 32484 35494: contig of 3000 bp in length
* 35495 35594: gap of unknown length
* 35595 40051: contig of 4457 bp in length
* 40052 40151: gap of unknown length
* 40152 41379: contig of 1228 bp in length
* 41380 41380: gap of unknown length
* 41380 44517: contig of 3038 bp in length
* 44518 44617: gap of unknown length
* 44618 45564: contig of 947 bp in length
* 45565 45664: gap of unknown length
* 45665 49440: contig of 3776 bp in length
* 49441 49540: gap of unknown length
* 49541 52698: contig of 3158 bp in length
* 52699 52798: gap of unknown length
* 52799 55964: contig of 3166 bp in length
* 55965 56064: gap of unknown length
* 56065 59236: contig of 3172 bp in length
* 59237 59336: gap of unknown length
* 59337 70052: contig of 10716 bp in length
* 70053 70152: gap of unknown length
* 70153 84603: contig of 14451 bp in length
* 84604 84703: gap of unknown length
* 84704 86418: contig of 1715 bp in length
* 86419 86518: gap of unknown length
* 86519 88069: contig of 1551 bp in length
* 88070 88169: gap of unknown length
* 88170 89661: contig of 1492 bp in length
* 89662 89761: gap of unknown length
* 89762 91357: contig of 1596 bp in length
* 91358 91457: gap of unknown length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 176100: gap of unknown length
* 176101 176200: contig of 6142 bp in length
* 176201 183079: contig of 6879 bp in length
* 183080 184752: gap of unknown length
* 184753 183180: contig of 1573 bp in length.
*
*
* location/Qualifiers
* 1. 184752
* /organism="Zea mays"
* /mol_type="genomic DNA"
* /strain="B73"
* /db_xref="taxon:4577"
* /clone="ZMBB-C0216K08"
* 2211. 2310
* /estimated_length=unknown
* 9942. 10041

```

```

gap      /estimated_length=unknown
16746..16845
/estimated_length=unknown
gap      32395..32494
/estimated_length=unknown
gap      35495..35594
/estimated_length=unknown
gap      40052..40151
/estimated_length=unknown
gap      41380..41479
/estimated_length=unknown
gap      44518..44617
/estimated_length=unknown
gap      45565..45664
/estimated_length=unknown
gap      49441..49540
/estimated_length=unknown
gap      52699..52798
/estimated_length=unknown
gap      55965..56064
/estimated_length=unknown
gap      59237..59336
/estimated_length=unknown
gap      70053..70152
/estimated_length=unknown
gap      84604..84703
/estimated_length=unknown
gap      86419..86518
/estimated_length=unknown
gap      88070..88169
/estimated_length=unknown
gap      89662..89761
/estimated_length=unknown
gap      91358..91457
/estimated_length=unknown
gap      133728..133827
/estimated_length=unknown
gap      137469..137568
/estimated_length=unknown
gap      163844..163943
/estimated_length=unknown
gap      169859..169958
/estimated_length=unknown
gap      176101..176200
/estimated_length=unknown
gap      183080..183179
/estimated_length=unknown

```

ORIGIN

Alignment Scores:

```

Pred. No.:      4,41e-66      Length:      184752
Score:          1222.00      Matches:      361
Percent Similarity: 32.0%      Conservative: 116
Best Local Similarity: 24.2%      Mismatches: 183
Query Match:      30.1%      Indels:      834
DB:              14          Gaps:      17

```

US-10-600-070b-2 (1-801) x AC155597 (1-184752)

```

Qy      90  PheTyrGlnValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIleArgAlaPhe 109
      107759 TATGCTCAAGATCTCTGTGGGAGCCACATTTCTTAGCGATGCGATTCGAGGCGTTTC 107818
Db      110 GluAlaArgValSerIysProProGlnPheGlyPheSerAspAspAlaLeuIleSerArg 129
      107819 GAGTCGCGGATAGCTAGCCACCTCAGTATGGGTACAGCACAGAAAGCTCTGTGGCGCA 107878
Qy      130 ArgGlnIleLeuGlnAlaGlySerGlnThrLeuSerAsnProArgSerArgArgGlnTyr 149
      107879 CGGCAAAATGCTGCAGATTCGCCAGATGATCTCTCAAAACCAAGAGCTGCGCAGCAATAC 107938
Qy      150 AsnGlnIleLeuLeuAspAspGlnGluAlaThrValIleThrAspValProTyrAspIys 169
      150  AsnGlnIleLeuLeuAspAspGlnGluAlaThrValIleThrAspValProTyrAspIys 169

```

```

Db      107939 GACCGTCGCTTTCCGAGACCGTGATGCCGCACTCACATGATGTTCCTGGGATTAAG 107998
Qy      170 Val----- 170
Db      107999 GT-GAGGTTATGGACACGAACGTCGAACAATGCTATCTGTTTCTTGTAATCTGTA 108057
Qy      171 -----Pro 171
Db      108058 AATAGCCCTGGCGGTAACTTGTTCCTCTCTCCGAAACTTGCACAGAGTTCCA 108117
Qy      172 GlyAlaLeuCyValLeuGlnGlyGlyGlnThrGluIleValLeuArgValGlyGlu 191
      108118 GGTTGCTGTGTGTGCTTTCAGAGCGTGGGAGGCAACAACGTGTGATGCAACTGGAGAG 108177
Qy      192 AlLeuLeuIysGlnArgLeuProIysSerPheIysGlnAspValValLeuValMetAla 211
      108178 CACTTGCTTCAGACCGTCCACCTAAGCGGTTCAAGAGATGTGGCTTACGATGGCA 108237
Qy      212 LeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProAspPheIleThr 231
      108238 TTGGCTTATGTGACATATCAAGGATGCTATGGCAGCAAGCCCTCCAGATGTATCTGC 108297
Qy      232 GlyTyrGlnPheValGlnGluAlaLeuIysLeuGln----- 244
      108298 TGTGTGAGAGGTGCTTGAAGAGGCACTGAAGCTCTCGAGGTGAACTAATTGCTTTAAG 108357
Qy      244 ----- 244
      108358 TTGCAGATCCAAATATGCTTAAGGTGCTGTATACATACAGTTGTTTAACTTAATGTG 108417
Qy      244 ----- 244
      108418 ACATGTCACTGTAGGCGCAAAATATGTGTGCACTACATACATTTTGTATGTAACA 108477
Qy      244 ----- 244
      108478 TGCTCCAGTTTATATATATTTGTTGCTGTATACCGATTAACACTTGTTCCTTGT 108537
Qy      244 ----- 244
      108538 AAACAAATATCATTTGGGTAGACATTATGATTAAGTACTTTCATCTGCACTTAGATCT 108597
Qy      244 ----- 244
      108598 CAGTAGCCCTTGATCATGATATCAATTATAAAAGCCGTAACCATTTGCTACATGCGC 108657
Qy      244 ----- 244
      108658 ACAGCATTAAATCTGTTCACTGTAAACCAATCTGTATGAACATATTAGAACTGCT 108717
Qy      244 ----- 244
      108718 GTATAGAGATATGAATATGCTCTAGTTTTCCTCTCTCTCTATTTCTTAATAATTATC 108777
Qy      244 ----- 244
      108778 ATCCCATATTGCAAGCTGTACACAAAAAATGTCTGTGATTTTAACTGTGAGTTT 108837
Qy      244 ----- 244
      108838 ATTTTCACTTATTATAGATCAGCGTGTGAAGTTTGTAGTCTGTCTTGAATTAGAGA 108897
Qy      245 -----GluGlnGlyAlaSerSer 251
      108898 TATCTTGTAGACATACTAATGAAACATTTTACTTATAGAGATGTGTCAAGCAATCT 108957
Qy      251 uAlaProAspLeuArgAlaGlnIleAspGlnThrLeuGlnGluIleThrProArgTyrVal 271
      108958 TGCACCTGAACTGTGTCACAGATTAATGAACCTTTGGAGGAATTAACCTCTGTCTGT 109017
Qy      271 IleGlnIleLeuGlyLeuProLeuGlnIysAspAspTyrAlaIleIysArgLeuAsnGlyLe 291
      109018 ATTGAGCTTCTTGTCTCTCTACTGATGAATAAATTAACGCAAGAGGTCT 109077

```


[illegible]

QY	608	lAlaaspmetleuylsglulaseValylleleulalaglValAla1le1yle	628
Db	110029	AAATACCCCTTG-----AAGATATCTGCTGGTACACTGCTTCACCT	110073
QY	628	u1serleupheserGlnlys-----	635
Db	110074	ATTTCAGGTGTAAAGTTCMAAGTCTTGCTGTGAACAAGTCACTTCGCTTTAGGGG	110133
QY	635	-----	635
Db	110134	CGAGTATGGGTCTGTGGCGTGTGTGACTCCATTGATGTGATGAAGAACTGTATCAGATA	110193
QY	636	----LyrheteuylusSerSerSerSerPheGlnArglysaPmetValSerSerMetG1	654
Db	110194	GCATTACTTTCCTCCAAAGTCACTG-GGTTTT-----GCTTGATGCTCTTGACA	110240
QY	654	userAspValAlaThrile-----	660
Db	110241	AAAGCGGTAGCAATATTTCCACATAAATGTGTCCAGGGCCATAACTTCCTATTCT	110300
QY	660	-----	660
Db	110301	ATATGTAAATGCATCATATGAGTGCACACATAAATGCTCATAGAACCGTATTATTGG	110360
QY	661	-----G1yserva	663
Db	110361	TTAAACGTGCGCTGTGGCCAGGAAATTAATCTTCATCTTTGTTGTTGAAGTCCAGC	110420
QY	663	lArgAlaaspAspSerGlnAlaLeuProArgMetAspAlaArgThrAlaGluAsn1leVa	683
Db	110421	GGCAGATGAAGGCCACTGGAATCTCTTAGATGTGACCGGAAGTGGCTGTGAAGACTTGT	110480
QY	683	1serlystTropGlnylsilelysSerleuAlaPheGlyProAspHis-----	698
Db	110481	TCGCAGGTGGCAAGACATCAATGCAAGGCTTTGGGGCCAGAAACACATGTCAAGCGCAT	110540
QY	698	-----	698
Db	110541	GCMAAGGTGAGGTCCCAAAATTTCCATCTTATAGCTGGCTAGCGCTTTGTGGTTGG	110600
QY	699	-----Arg1leglu-----	701
Db	110601	TCCCAACAATATCCCAACACACACATGTGTCTGAGAGTTGAGAGGCACTTTGGCTGTGC	110660
QY	701	-----	701
Db	110661	GGTCTTTGAAAACTCGACCTGCGGGGGGTAAAGACAGCCCCGAGCATTTGATTAAAGAAA	110720
QY	701	-----	701
Db	110721	GACCTTCTCACACAGGTGAGAAAAACCCCGAACCCCTGCCCCACCCATACACAGCGCA	110780
QY	701	-----	701
Db	110781	TCGAAGCCCATGTGAGAAACGACACGACCGGGGCTGAGCCTTAGACTGTGCTTGGCGT	110840
QY	701	-----	701
Db	110841	GGGACAGACGAGAAGATTTTTTAAACACACAGCCTGAATTTGCGTCCGACGGGAGTCGAA	110900
QY	701	-----	701
Db	110901	CCCAGAGACTGAGAGTGTACTCAAAACCACTTAACCACTGACGCGCAATGTGSGGA	110960
QY	701	-----	701
Db	110961	CTTGCTTAAATGCTATGAATTAAGAAAAAGGCAACAAAATTTGTTAAAGCTTAAAG	111020
QY	701	-----	701
Db	111021	TGCTTCGTCTTCGAAAGCATTAATTTCTGTGTGGATATATAGAGTTTCAGACGAAGTTAT	111080
QY	701	-----	701


```

Db      111081 GAAGAGCTACCTTCATCAGACAACAATAATGAGAGATTCTATATAAATATAGA 111140
QY      701 -----701
Db      111141 AATAACATMAAACCTTAATATATATATATTAATCATCTTCTATTTATATATATGAAAAA 111200
QY      701 -----701
Db      111201 TAGAATAATATACAAATTAACAATATGATCTTGGCTTGAAAGATGTAAAAAGTAAACAGC 111260
QY      701 -----701
Db      111261 TAGATCAAAAAGCAATGCAAGTCAAGGTGAACAGTACGGAGCATTTGTCATCTATTTA 111320
QY      701 -----701
Db      111321 TAGGACGAGACGACGCCATGTAAATACTTATGCTTACCTTTACCTGCTAATAACT 111380
QY      701 -----701
Db      111381 CTATAGTATATCATCGGGGTCTAATAATAGTCTTTCTCTTAAAGTCGGTTCAATTTCTG 111440
QY      701 -----701
Db      111441 CTATACGCGGAAGCTCCCTGCGCATAGCTTGGCTCTGGCCATCTTCATTTCTTT 111500
QY      701 -----701
Db      111501 GTGCTTCTTCACACTGTGTGTTTGACGTAAGTCCGAAGTACCTGTTCAATGATCATACT 111560
QY      701 -----701
Db      111561 CCAAAAATTGTTAAATCAATGTTTGAAGACCTTGGAAGCGCAAGCCCCCAACAA 111620
QY      701 -----701
Db      111621 TCACGCAATTACATTGAAGAAGATTTTGAAGCATTTGAGATTGACGACCAACCACTCG 111680
QY      701 -----701
Db      111681 TTCAATATAGTGAAGTGTGCTTCTTATGAAACCCCTGCTGTATGACTAAATCTCAG 111740
QY      702 -----MetLeuProGluValLeuAspGlyArgMe 711
Db      111741 TGCACCTGTACATGACGTGCTTCCAAATTCATTCGCGCAGGTCTCTCGCGCAACAT 111800
QY      711 tleuylsetrtrthraparaglaaiaagiutthralaglnleuglyleuvaltyraeply 731
Db      111801 GCTGAAGCTGTGGCGGACCGAGCTGCAAGATCGAGCGCCACGCGTGTCTGGAGATA 111860
QY      731 rthleuLeuLysleuSerValAspSerValThrValSerAlaAspGlyThrArgAlaLe 751
Db      111861 CGGCTCTCTGGCGTGAAGTCAACAGCATCACCGTCTCGGTGAACGCGCGCGGAGC 111920
QY      751 uvalgluaiaLthleuglnleuSerAlaCybleuSerAspLeuValHisProGluAsnAs 771
Db      111921 GGTGAGGCGCACATCGAGAGGTGGCGCGCTTACGAGGTGACCGACCTTAAGAACGA 111980
QY      771 naIatThrAspValArgThrTyThrThrArgTyGluValPheTrpSerLys---SergI 790
Db      111981 CGACCGCTAGACACCAAGTACACCGCCGCTAGAGATACCTATCTCCAGGCGACGAGG 112040
QY      790 YTPlyValIethrGluGlySerValLeuAlaSer 801
Db      112041 GTGAGAGATCACCGAAGCGCAGTCTCAAGTCTG 112074

```

```

CR954204_1 100001 210000
CR954204_2 200001 310000
CR954204_3 300001 410000
CR954204_4 400001 510000
CR954204_5 500001 610000
CR954204_6 600001 710000
CR954204_7 700001 810000
CR954204_8 800001 884654
CR954204_8 884654 bp DNA linear HTG 05-JUL-2005
LOCUS CR954204
DEFINITION Ostreococcus tauri strain OTTH0595, *** SEQUENCING IN PROGRESS ***
ACCESSION CR954204
VERSION CR954204.2 GI:70062289
KEYWORDS HTG; HTG; PHASE2; green alga; whole genome shotgun.
SOURCE Ostreococcus tauri
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;
Mamiellaceae; Ostreococcus.
REFERENCE 1 (bases 1 to 884654)
Dereille,E., Ferraz,C., Rombauts,S., Rouze,P., Morden,A.Z.,
Partensky,F., Degroove,S., Echeynie,S., Cooke,R., Robbens,S.,
Wuylt,J., Saeys,Y., Jabbari,K., Bowler,C., Ball,S., Ral,J.P.,
Bouget,F.Y., Pigneanu,G., De Baets,B., Picard,A., Delseny,M.,
Demaille,J., Van de Peer,Y. and Moreau,H.
Genome analysis of the smallest free-living eukaryote Ostreococcus
tauri unveils unique genome heterogeneity
Nature (2005) In press
2 (bases 1 to 884654)
Rombauts,S., Dereille,E., Ferraz,C., Van de Peer,Y. and Moreau,H.
Direct Submision
Submitted (30-APR-2005)
On Jul 7, 2005 this sequence version replaced gi:62990096.
COMMENT Contact: Moreau H
Laboratoire Arago
CNRs UMR 7628
BP 44, Avenue Fontaule, 66651 Banyuls sur mer, France Tel:
(33)468887309
Fax: (33)468887398
Email: h.moreau@obs-banyuls.fr
----- sequencing/assembly
Institut de Genetique Humaine, UPR CNRS 1142
34396 Montpellier Cedex 5, France
----- annotation
Department of Plant Systems Biology,
Planters Interniversity Institute for Biotechnology (VIB), Ghent
University, B-9052 Ghent, Belgium
Web site: http://bioinformatics.psb.ugent.be/genomes.php

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 884654: contig of 884654 bp in length.
Location/Qualifiers
source
1. 884654
/organism="Ostreococcus tauri"
/mol_type="genomic DNA"
/strain="OTTH0595"
/db_xref="taxon:70448"
/clone_lib="Ostreococcus tauri genomic shotgun library"
/notes="assembly version 050606"

```

ORIGIN

```

RESULT 11
CR954204_0/c
WPCOMMENT
Sequence split into 9 fragments LOCUS CR954204 Accession CR954204
Fragment Name Begin End
CR954204_0 1 110000

```

```

Alignment Scores:
Pred. No.: 2,99e-33 Length: 110000
Score: 702.00 Matches: 220
Percent Similarity: 45.5% Conservative: 116
Best Local Similarity: 29.8% Mismatches: 291
Query Match: 17.3% Indels: 112

```

DB: 14 Gaps: 25
 US-10-600-070b-2 (1-801) x CR954204_0 (1-110000)
 QY 78 AapRyPRGluIuRghIsvAlProIIePhrPhyrgIuValIeugIyAlaGln 97
 25764 GACGCCCGGAGGCGACGCGCGCTCCCGCTTCGACTATCAAACTTGGGATGGTCG 25705
 QY 98 ThHIsPheIeuThr---AapGlyIleAryAlaPheGluIuIaArGValSerTyPro 116
 25704 CCGCGCGAAACGACCGGAAAGCGGCTCCACGACGCGCTCCGCGGATGAGACGCCG 25645
 QY 117 ProGlnPheGlyPheSerAapAlaIeuIleSerAryrgIuIleuGlnIaIa 136
 25644 CTTCGCCAGGCTTACTCCGCAACTGCTCGAACACAGCTGCTCATCGAACAGCC 25585
 QY 137 CyseGluThrIeuSerAenProArGSerAryrgIuTyAenGluTyLeuIeuAapA 156
 25584 GTGGCGGTGCTCAAGATGAAGAGGCTCGCGACAGACGATGAG-----GAC 25537
 QY 157 GluGluIuAlaThrValIleThrAapValProTrAapRyAValProGlyAlaIeuCyA 176
 25536 ATTCGCGAGGTGTGCTCACTCCGCTGGAAGCGCATGCGCGCGCGCTGTGTCTTA 25477
 QY 177 LeuGlnIuGlyIuGlyIuThrGluIleValIeuAryValGlyIuAlaIeuLeuTyGlu 196
 25476 CTTCAGAGGCGCGCGGATCGAATCCGTGATGAGTTGAGCCGCTCGCGCAAGTCG 25417
 QY 197 ArgLeuProIySerPheIyGlnAapValValIeuValMetAlaIuAlaPheIeuA 216
 25416 GTAGCGGACGACGACCGCAAGAGATGCGCGCTCACCGCGCTCGCTGTGTGAG 25357
 QY 217 ValSerAryAapAlaMetAlaIeuAapProProAapPheIleThrGlyTyrgIuPhA 236
 25356 TTCGCGCAATGCGGCTCATGCGAATCCGCTCGCATCTCCGAGGTGCGAACTGCTC 25297
 QY 237 GluGluIuAlaIeuTyIleuIeugIuGlnIuGlyAlaSerSerIeuAlaProAapLeuA 256
 25296 GAGATGGGATCGAAGACGCTCTCGAAGTGGC---CGGTCTTCGCGCGCGAAGTTCGT 25240
 QY 257 AlaGlnIleAapGluThrIleuGluIuIleThrProAryTyValIeugIuLeuTyGly 276
 25239 CGCAACATCTCGGCTTATATCAACAGCTCGCTCGGGCTATGAGTCAATTGTGTGCG 25180
 QY 277 LeuProIeuGlyAapAryTyAlaAlaIyAryLeuAenGlyIeuSerGlyValAryAsn 296
 25179 ATGCGCGTCTC---GAGCGCGAGTCCGAGCGCGCTCGCGCGCTTCGAGC 25126
 QY 297 IleuTyPserValGlyIyAlaIaSerAlaIeuValGlyIyLeuThrAryGlu 316
 25125 CTCTTTTGAACAGACCGCTCGCAACAGCTCGAA-----CAGCGCGCG 25081
 QY 317 LysPheMetAenGluAlaPheLeuAryMetThrAlaIaGluIuValAapLeuPheVal 336
 25080 GCGTTCATGGAACAGACGAAATTACTCAACGCGCAAGAGCAGTCCGCACTCTTCGTC 25021
 QY 337 AlaThrProSerAenIleProAlaGluSerPheGluValTyrgIuValAlaIeuAlaIeu 356
 25020 GAGCGTCCGATTAATCCCGCTTGACTCGACGAAATGTAACAGACGCGCTCGCCAC 24961
 QY 357 ValAlaGlnIaPheIleGlyIyLysPheProHIsIeuIeugIuAapAlaAapTyGlnPhe 376
 24960 GTGCTCCGCGGTGTCATCGAATGCCAATGCTTCGTCGACGCGGAGAAATCTTG 24901
 QY 377 GlnGlnIeugIuGlnIaIyValMetAlaMetGluIleProAlaMetLeuTyAapThr 396
 24900 GAGCAATTAACGCTCGCTGAGCTCCCTGCGAATCAAGCGCTTC-----24853
 QY 397 ArgAenAenTrpGluIleAapPheGlyIeugIuAryGlyIeuCyAlaIeuLeuIleGly 416
 24852 -----TCCGACGTGCGCGTCAAGCGCGCGGTGTGCAAAATTTCTTGTGT 24808
 QY 417 LysValaIeuGluCyAryMetTrpLeuGlyIeuAapSerGluAapSerGlnTyAryAsn 436

DB 24807 CGCATTTGAGAAAGCCGATGACTCTGGGTTCATCAACGATACCGCT-----GAT 24757
 QY 437 ProAlaIleValGluPheValIeugIuAenSerAenAryAapAapAapLeu-----455
 DB 24756 CCGAGCTGTGTCAATACATC-----GAGATCGCTCCCGCAATGGAATCTCGCG 24706
 QY 456 ProGlyIeuCyIyIeIeuIeugIuThrTrpLeuAlaGlyValIaPheProAryPheAry 475
 DB 24705 GAGGTATGTGTGTCATGCGGATCGAGTCAAGTGTGATCGACGTGCTCCGCTTCGCCG 24646
 QY 476 AapThrLysAapIyIyIyPhe---LysIeuGlyAapTyTyThrAapAapProMetValIeu 494
 DB 24645 GGGACAGACGACCAACCAACCGCCACCATCGAGATGTGTTCAACGCGCAAGCTTCGA 24586
 QY 495 SerTyIeuGluIyValGluValIaGlnGlySerProIeuAlaIaIaIaThrMet 514
 DB 24585 GGAATTTGTAGCCGATTTGAT-----24565
 QY 515 AlaAryIleGlyAlaGluIuIleValIyAlaSerAlaMetGlnAlaIeugIuTyValPhe 534
 DB 24564 -----TCGATGCGCGCTCGTTGCA-----24544
 QY 535 ProSerAryTyThrAapAryAenSerAlaGluProIyAapValGlnIuThrValPhe 554
 DB 24543 -----ATCCAGCGCGGATCGAA 24526
 QY 555 SerValAap---ProValGlyAenAenValIyAryAapGlyGluProGlyValPheIle 573
 DB 24525 TCGGCAAGAGGACCGCTGCTCGCGCTCGGCTCG-----ATT 24487
 QY 574 AlaGluAlaValAryProSerGluAenPheGluThrAenAapTyTyAlaIaIaIaGly 593
 DB 24486 GCGACGCTTTTCTTCGACTCGCGCG-----GCGCTGAGTTGTGT 24445
 QY 594 ValSerGluSerValAapGluThrThrValGluMetSerValAlaAapMetLeuTy 613
 DB 24444 ATCTCCGAT-----ACCAACGACGAGCGCTTCCT-----24415
 QY 614 GluAlaSerValIyIleIeuAlaIaGlyValAlaIaIeGlyLeuIleSer-----630
 DB 24414 GTGGCAAGATTTGTCTCTC---GCCGAGTAGCGCTTGTTCACACGAAATGAGCC 24358
 QY 631 ---LeuPheSerGlnIyTyThrPheIeuTyIySerSerSerPheGlnIyTyAapMet 649
 DB 24357 GGTCTGCTCGCGGAGAGATACCGCGAGCGCGCGGTTCCTCGTTGCAAGCG 24298
 QY 650 ValSerSerMetGluSerAapValAlaThrIleGlySerValAryAlaAapSerGlu 669
 DB 24297 GTCGCGGCGCGCTGTCAACGTCAAGTTGTGACCGCGACACGACAAACGACGCGC 24238
 QY 670 AlaIeuPro-----AryMetAapAlaAryThrAlaGluAenIleValSerTyTrpGln 687
 DB 24237 CCGGCGCGCTGTACCAACATGCAAGGTAACCGGGAAGATAGTGGCAATGGCAG 24178
 QY 688 LysIleIySerIeuAlaPheGlyTyProAryAlaIaGluThrAlaGlnIeugIyIeu 707
 DB 24177 ACGGCCAAGCGAGGCTTTGGGGCAAGCCACACATGCGCAGCTGAAAGAAATCTTA 24118
 QY 708 AapGlyAryMetLeuTyIleTrpThrAapAryAlaIaGluThrAlaGlnIeugIyIeu 727
 DB 24117 GAGGGGCGCATGTTGCAAGATGCAAAACCGCGGAGGACGTCAAGCGCGCAGATGCG 24058
 QY 728 ValTyAapTyThrIeuIeuTyIyIeuSerValAapSerValThrValSerAlaAapGly 747
 DB 24057 GCGTGGAGATACCAACTGAATGAATGATATGACAC-----GTTCAAGCGCTTGA 24004
 QY 748 Thr---AryAlaIeuValGluAlaThrIeugIuIySerAlaCyIeuSerAapLeuVal 766
 DB 24003 ACTGAGAAATTTTGTGAGACAGCGCTCACTGAGGTGCGGGTGTGAAGAT-----23950
 QY 767 HisProIuAenAenAlaThrAapVal-----AryThrTyThrThrAryTyrgIuVal 784

Db 23949 CACGCGAAGACCAACGATGACGTGTACGAGACGATCGAGCCAACTAGACGCTT 23890
QY 785 -----PheTpsrSerGlyTTrpValLeuThrGluGlySerValLeu 799
Db 23889 AAAAGTGTCTGACGAGCTGAAAGAGGAGTGTGGAGGAAATGTGATT 23833

RESULT 12
BA000019_32/c
MPCOMMENT

Sequence split into 65 fragments	LOCUS	Accession
Fragment Name	Begin	End
BA000019_00	1	110000
BA000019_01	100001	210000
BA000019_02	200001	310000
BA000019_03	300001	410000
BA000019_04	400001	510000
BA000019_05	500001	610000
BA000019_06	600001	710000
BA000019_07	700001	810000
BA000019_08	800001	910000
BA000019_09	900001	1010000
BA000019_10	1000001	1110000
BA000019_11	1100001	1210000
BA000019_12	1200001	1310000
BA000019_13	1300001	1410000
BA000019_14	1400001	1510000
BA000019_15	1500001	1610000
BA000019_16	1600001	1710000
BA000019_17	1700001	1810000
BA000019_18	1800001	1910000
BA000019_19	1900001	2010000
BA000019_20	2000001	2110000
BA000019_21	2100001	2210000
BA000019_22	2200001	2310000
BA000019_23	2300001	2410000
BA000019_24	2400001	2510000
BA000019_25	2500001	2610000
BA000019_26	2600001	2710000
BA000019_27	2700001	2810000
BA000019_28	2800001	2910000
BA000019_29	2900001	3010000
BA000019_30	3000001	3110000
BA000019_31	3100001	3210000
BA000019_32	3200001	3310000
BA000019_33	3300001	3410000
BA000019_34	3400001	3510000
BA000019_35	3500001	3610000
BA000019_36	3600001	3710000
BA000019_37	3700001	3810000
BA000019_38	3800001	3910000
BA000019_39	3900001	4010000
BA000019_40	4000001	4110000
BA000019_41	4100001	4210000
BA000019_42	4200001	4310000
BA000019_43	4300001	4410000
BA000019_44	4400001	4510000
BA000019_45	4500001	4610000
BA000019_46	4600001	4710000
BA000019_47	4700001	4810000
BA000019_48	4800001	4910000
BA000019_49	4900001	5010000
BA000019_50	5000001	5110000
BA000019_51	5100001	5210000
BA000019_52	5200001	5310000
BA000019_53	5300001	5410000
BA000019_54	5400001	5510000
BA000019_55	5500001	5610000
BA000019_56	5600001	5710000
BA000019_57	5700001	5810000
BA000019_58	5800001	5910000
BA000019_59	5900001	6010000
BA000019_60	6000001	6110000
BA000019_61	6100001	6210000

BA000019_62 6200001 6310000
BA000019_63 6300001 6410000
BA000019_64 6400001 6413771
Continuation (33 of 65) of BA000019 From base 3200001 (BA000019 Nostoc sp. PCC 7120 DNA

Alignment Scores:

Pred. No.:	3,38e-20	Length:	110000
Score:	497.00	Matches:	207
Percent Similarity:	41.0%	Conservative:	139
Best Local Similarity:	24.6%	Mismatches:	301
Query Match:	12.2%	Indels:	196
DB:	1	Gaps:	33

US-10-600-070b-2 (1-801) x BA000019_32 (1-110000)

QY 84 ValProIleProIlePhePheTyrGluValLeuGluValAlaGluThrIlePheLeuThrAsp 103
Db 102796 GTGCGAATTCGGCTAGATTATTCACGGAATTTTGGGCTACCGTTAGCGGCAAGTGAATGAA 102737
QY 104 GlyIleArgArgAlaPheGluAlaArgValSerIleProGluThrIlePheSerAsp 123
Db 102736 CAATGCGACAGCAACATACAGCCATTCGATTCGCAATTCGCGAGGAGATTCCTCA 102677
QY 124 AspAlaLeuIleSerArgArgGluIleLeuGluAlaAlaCysGluThrIleSerAsp 143
Db 102676 GCGCAATTCCTCCCTTAACACTTATAGAGAAGCTTACGTTTATCATGATCCA 102617
QY 144 ArgSerArgArgGluTyr-----AsnGluGlyLeuLeuAspAsp 149
Db 102616 AAGGAACGACAGATTATGACGAGTGTATCTTGTCTACGCGCTACGACCAACGCAACGCG 102557
QY 150 -----AsnGluGlyLeuLeuAspAsp 156
Db 102556 GCTACACCAAGATGCGACGTGAAAATCGTGGGACAGCAACATGTCATTTGATGTC 102497
QY 157 GluGluAlaThrValIleThrAspValProTyrAspValProGluValLeuCysVal 176
Db 102496 CAAGCTGAGCATC-----GAGTTTCCTCGAGAAATTAATGTCCTTTATTAAT 102443
QY 177 LeuGluGluGlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeu----- 194
Db 102442 TTGCAAGATTGGGAGATGATGAACTGATCAAGTTAGTGTGATTAATTAAT 102383
QY 195 -----LysGluArgLeuProLysSerPhe----- 202
Db 102382 CAAAACGACAGATCCACAGAAATGCAATCATCGACGCTGAAAGAAATTTCTCGAT 102323
QY 203 -----LysGluAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArg 219
Db 102322 AGTTCTGAACGTCAGATATTTCTCTTGACTGTTCCTTGCTCATTTAGAGGCGG 102283
QY 220 Asp-----AlaMetAlaLeuAspProAspPhe 229
Db 102262 GAACATGSCAACAGGCCCATATGAAAACGCTGTTCTCTTAGAG----- 102215
QY 230 IleThrGlyTyrGluPheValGluGluAlaLeuLysLeuGluGluGlyAlaSer 249
Db 102214 ---ACTGGGCAAGAA-----GTCCTGTTTAGTGAAGC----- 102185
QY 250 SerLeuAlaProAspLeuArgAlaGluIleAspGluThrLeuGluGluThrProArg 269
Db 102184 ---ATCTTCCCAAGCGCTCCAGGCAAGAAATTCAGCGTATCTTACAAATTAAGCCCTTAT 102128
QY 270 TyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyrAlaAlaLysArgLeuAsn 289
Db 102127 ACAATTTTGAATTAATCTTACCTTACCC-----CAGAGAAAACCATGACGCAACCA 102074
QY 290 GlyLeuSerGlyValArgAsnIleLeuTpsrValGlyGly---GlyGlyAlaSerAla 308
Db 102073 GGGCTGATCTATTCAAAGCATTCCTTAGACATCCGCTGGGATGATGATGATGATGAT 102014
QY 309 LeuValGlyGlyLeuThrArgGlu-----LysPheMetAsnGluAlaPheLeuArg 325

Db 102013 GATCATCAGGCTTAAACATTGATGACTTCTCCGATTGATCCAGCAATTAAGCCACAC 101954
 QY 326 MetThrAlaAlaGlnGlnValAlaSerPheValAlaThrProSerAsnLeuProAlaGlu 345
 Db 101953 TTAAACGCGCTGAAACAACTAAAGTTGTGATGCT---GAAAGCAAGCCCTTGTG--- 101900
 QY 346 SerPheGluValAlaValAlaAlaLeuAlaValAlaGlnAlaPheLeuLeuGlyLeuVal 365
 Db 101899 GCTGTGCTCATACCTTACCTGCTTTTGTCTTCATGCGCAGAGATTTACCCCAAGCCAG 101840
 QY 366 ProHISLeuLeuGlnAlaSerAlaSerGlnPheGlnGlnLeuGlnAlaValValMet 385
 Db 101839 CCCGCTTTATTCGTATGTCAGCAAGCAAAATTCGTATGCTGCTTAAGCGCGCA--- 101786
 QY 386 AlaMetGluIleProAlaMetLeuIleTyraSerThrArgAsnAsnTrpGluLeuSerPheGly 405
 Db 101785 -----GATGTGCAAT 101777
 QY 406 LeuGluArgGlyLeuGlyCysAlaLeuLeuIleGlyLeuValAlaSerGluCysArgMetTrpLeu 425
 Db 101776 TTAGACGACCTCCTGTGTGCGCTTATTACTAGGCGAACTGAGAAAGCCAGCGAGTTTA 101717
 QY 426 GlyLeuAspSerGlnAspSerGlnTyraArgAsnProAlaIleValGluPheValLeuGlu 445
 Db 101716 GAACTGAGCCAAAGA-----TACGAAGCTTTAGCCTTAAT---CGA 101678
 QY 446 AsnSerAsnArgAspAsnAsnAspAsnAspLeuProGlyLeuGlySerLeuLeuGluThrTrp 465
 Db 101677 GAAAAATCTCAAGATTCACCCGATTTTACGTCAGGTTTGCTTATATAGCCGAACAATGG 101618
 QY 466 LeuAlaGlyValAlaPheProArgPheArgAspThrIleAspIleValSerPheLeuGly 485
 Db 101617 CTGCAGAAATGAAGATTTTCCCTCCATTTCCGCGATTTGTCCAGACAGCAAGCTTCCCTGAAA 101558
 QY 486 AspTyrTyrAspAspProMetValLeuSerTyrLeuGlnArgVal----- 500
 Db 101557 GATTACTTGTCTATATCAACAAGATGACGATTTTGAAGCCTTGCCCAACGACGCGGAA 101498
 QY 501 -----GluValAlaGlnGlySerPro 507
 Db 101497 ACCACTAATGAATGGCTGTAAATTAACCCGCAATCGTTTCTCAACGAGGCGCAATTC 101438
 QY 508 LeuAlaAlaAlaAlaThrMetAlaArg-----IleGlyValaGlnHis-----ValIys 523
 Db 101437 TACTCTGGAGGAACGCCAGTCCGCAACGTCCTCCGTAAGGAGCGAAACAGCCAGAGAA 101378
 QY 524 AlaSerAlaMetGlnAlaLeuGlnValPheProSer-----ArgTyrThrAsp 540
 Db 101377 GCGTCCACAAAGCAAGTTCGCCCAAGTCATCCATCAGAGTAATCGCAGTTTCAT 101318
 QY 541 ArgAsnSerAlaGluProIlyAspValGlnGluThrValPheSerValaProValGly 560
 Db 101317 CAAAAACGAACCCCTGATCCCGAATTACGAAAAATCAAAACACAGAGACAGAGTCT 101258
 QY 561 AsnAsnValGlyArgAspGly----- 568
 Db 101257 TCAAAATTTTACAACTGCTAGAGAAATATATTCAGACACAGATGCTTACCTGCAATTA 101198
 QY 569 ---ProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
 Db 101197 CCACCAAGATCCCTGTAGAACCGCGCAGACACT----- 101162
 QY 588 TyrAlaIleArgAlaGlyValSer-----GluSerSerValaSerGluThrThrVal 604
 Db 101161 -----GTTCAAGCGGGGTAGAGTATATACCAATCGACCCCTCCACGCAAACTCCT 101108
 QY 604 ----- 604
 Db 101107 AAACGAGAGAGCGCAAGAAACCAAGCAGGTGTCAACAGAGACACAGATTCATCAG 101048
 QY 605 -----GluMetSerValaIleAspMetLeuGlyGluAlaSerValIleLeuAlaAla 622
 Db 101047 CAACGCCAACCTCCTACCTAGACCTTAGGCGGAAACAGATTAATCTTGATATGTTTG 100988

QY 623 GlyValAlaIleGlyLeuIleSerLeuPhe-----SerGlnIleTyr---Phe 637
 Db 100987 GGT---TCTTGGGCGGATATATGTTCTGCTGATAGTCAACAGACTTTGGGTGG 100931
 QY 638 LeuIlySer-----SerSerPheGlnAlaGlyIleAspMetValSerMet 653
 Db 100930 TTAAAGAAATGATATCTTCCACAGCACTTTTACAGGTGACCAATTAATCGATTGAT 100871
 QY 654 GluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeuProArg 673
 Db 100870 AGTCAACCACTTTAGAGATTTCTGACAAATAATGCCAGATACATCCCCAGAGTGACT 100811
 QY 674 MetAspAlaArgThrAlaGluAsnIleValSerIleTrpGlnIleIleIysSerLeuAla 693
 Db 100810 CTCACAGAAAGAAACCGCAAGAAATTAATGAAAATTTGGTGGCTACCAAGACTAGTGT 100751
 QY 694 PheGlyProAspPheArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuVal 713
 Db 100750 TTAGCGGCTGAACATTAATTAAGAGTTTAAACGAGATTTTAACTGTTCAGCGTTACT 100691
 QY 714 IleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeu 733
 Db 100690 CAATGCGGCTAATTTGCTTGCAGATTAAGACAGACATCTGCATCCAGAAATCACT--- 100634
 QY 734 LeuIysLeuSerValaSerSerValaSerValaSer---AlaAspGlyThrArgAlaLeuVal 752
 Db 100633 CATAGTGCAAGTATGACTGATCCATCAATGATGATGATGATCCCAATGTCGAAGTGTG 100574
 QY 753 GluAlaThrLeuGluGluSerAlaCysIleuSerAspLeuValHis-----ProGluAsn 770
 Db 100573 GGGGCTACAGTCAGAGAG-----TTAACCCAAATTTATGAGAAATGGGCAAAAGG 100523
 QY 771 AsnAlaThrAspValaArgThrTyrThrThrArgTyrGluValPheThrSerIysSerGly 790
 Db 100522 AAGTCTTCAGAGAAAGA---TTAGCTGTACGCTATGAATGATTTGACAAAGTATAT 100466
 QY 791 TrpIleVal 793
 Db 100465 TGGCGGATT 100457

RESULT 13
 BA000019_33/c
 WPCOMMENT
 Sequence split into 65 fragments LOCUS BA000019 Accession BA000019
 Fragment Name Begin End
 BA000019_00 1 110000
 BA000019_01 100001 210000
 BA000019_02 200001 310000
 BA000019_03 300001 410000
 BA000019_04 400001 510000
 BA000019_05 500001 610000
 BA000019_06 600001 710000
 BA000019_07 700001 810000
 BA000019_08 800001 910000
 BA000019_09 900001 1010000
 BA000019_10 1000001 1110000
 BA000019_11 1100001 1210000
 BA000019_12 1200001 1310000
 BA000019_13 1300001 1410000
 BA000019_14 1400001 1510000
 BA000019_15 1500001 1610000
 BA000019_16 1600001 1710000
 BA000019_17 1700001 1810000
 BA000019_18 1800001 1910000
 BA000019_19 1900001 2010000
 BA000019_20 2000001 2110000
 BA000019_21 2100001 2210000
 BA000019_22 2200001 2310000
 BA000019_23 2300001 2410000
 BA000019_24 2400001 2510000
 BA000019_25 2500001 2610000
 BA000019_26 2600001 2710000


```

QY 501 -----GluValValGlnGlySerPro 507
Db 1497 ACCACTAATGATGGGCTGTAATTAAACCGCAATCGTTTCTCAACCCGAGGGCAATTCT 1438
QY 508 LeuAlaAlaAlaThrMetAlaArg-----IleGlyValGlnHis-----ValLys 523
Db 1437 TACTCTGAGAGAAACGCGACATGCGCAAACTGCGTAGGGAGGCGAAACAGCCAGAGAGAA 1378
QY 524 AlaSerAlaMetGlnAlaLeuGlnLysValPheProSer-----ArgTyrThrAsp 540
Db 1377 GCGTCCCAAGACCAAGTCTCCCAACAGATCATCATCAAGAGTAATTCAGCGATTTCAT 1318
QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
Db 1317 CAATAACGAACCCCTGATCCCGAATTACGAAACATCAAAACCAAGAACCAAGAGCTC 1258
QY 561 AsnAsnValGlyArgAspGlyLys----- 568
Db 1257 TCAAAATTTTCAACTGCTAGAGAAATATATCGAACCAAGATGCTTACACTGACAAATTAT 1198
QY 569 ---ProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
Db 1197 CCACCAAGATCCCTGTAGAACCGCGCCAGCAACT----- 1162
QY 588 TyrAlaIleArgAlaGlyValSer-----GluSerSerValAspGluThrThrVal 604
Db 1161 -----GTTACGCCGGGGGTAGTGTTATACCAATCGACCCCTCCACGGCAAACTCCT 1108
QY 604 ----- 604
Db 1107 AAACGACAGACGCAAGAACCAAGCAGATTGTCAACAGAGACACAGTATTTCATCAG 1048
QY 605 -----GluMetSerValAlaAspMetLeuLysGlnAlaSerValLysIleLeuAlaAla 622
Db 1047 CAACGCCAACCCCTACCTAGCCTTAGCGCCAAACAAAGATTCTTGGAATGTTTG 988
QY 623 GlyValAlaIleGlyLeuIleSerLeuPhe-----SerGlnLysTyr---Phe 637
Db 987 GGT---TCTTGGGTGGGATATATTTCTTGGCTGATCTCAACGACTTTTGGGCTG 931
QY 638 LeuLysSer-----SerSerSerPheGlnArgLysAspMetValSerSerMet 653
Db 930 TTAAAGATGTATTTCTTCCAGACACCATTTTAAAGTAGAGCAATATGATTCAGATT 871
QY 654 GluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeuProArg 673
Db 870 AGTCAACACACTTAGAGATTCTCGACAAATAATGCCAGATACATCCCAAGAGTGAGT 811
QY 674 MetAspAlaArgThrAlaGluAlaAsnIleValSerLysTyrGlnLysIleLysSerLeuAla 693
Db 810 CTCACAGAGAAACGGCAGAGAAATTAATTGAAATTTGGTCTACCAAGCTAGTGT 751
QY 694 PheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuLys 713
Db 750 TTAGCGCGCTGAACATAAAATTGAGAGTTTAAACGAGATTTTAACTGGTTCAGCGTTATCT 691
QY 714 IleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeu 733
Db 690 CAATGGCGGCTAATTCCTTGCAGAGATMAAGACAGACATCGTCATCGCAATTCAGT--- 634
QY 734 LeuLysLeuSerValAspSerValThrValSer---AlaAspGlyThrArgAlaLeuVal 752
Db 633 CATAGTGTCAAGTAGATCTCATCATCAGTAATTCGACATAGATCCCAATCTGTCAAGTGTG 574
QY 753 GluAlaThrLeuGlnGluSerAlaCylLeuSerAspLeuValHis-----ProGluAsn 770
Db 573 GGGGCTACAGCTCAGAGG-----TTAACCAATTTTATGATGAGATGGCAAAAGGG 523
QY 771 AsnAlaThrAspValAlaGlnTyrTyrThrThrArgTyrGluValPheTyrSerLysSerGly 790
Db 522 AAGCTTGTGACGAAAGA---TTACGTGTACGCTATGATTTGATTCGACAAAGATGATATT 466
QY 791 TrpLysIle 793

```

```

Db 465 TGCGGATT 457
RESULT 14
BA000045_36
WPCOMMENT
Sequence split into 47 fragments
Fragment Name Begin End
BA000045_00 1 11000
BA000045_01 10001 21000
BA000045_02 20001 31000
BA000045_03 30001 41000
BA000045_04 40001 51000
BA000045_05 50001 61000
BA000045_06 60001 71000
BA000045_07 70001 81000
BA000045_08 80001 91000
BA000045_09 90001 101000
BA000045_10 100001 111000
BA000045_11 110001 121000
BA000045_12 120001 131000
BA000045_13 130001 141000
BA000045_14 140001 151000
BA000045_15 150001 161000
BA000045_16 160001 171000
BA000045_17 170001 181000
BA000045_18 180001 191000
BA000045_19 190001 201000
BA000045_20 200001 211000
BA000045_21 210001 221000
BA000045_22 220001 231000
BA000045_23 230001 241000
BA000045_24 240001 251000
BA000045_25 250001 261000
BA000045_26 260001 271000
BA000045_27 270001 281000
BA000045_28 280001 291000
BA000045_29 290001 301000
BA000045_30 300001 311000
BA000045_31 310001 321000
BA000045_32 320001 331000
BA000045_33 330001 341000
BA000045_34 340001 351000
BA000045_35 350001 361000
BA000045_36 360001 371000
BA000045_37 370001 381000
BA000045_38 380001 391000
BA000045_39 390001 401000
BA000045_40 400001 411000
BA000045_41 410001 421000
BA000045_42 420001 431000
BA000045_43 430001 441000
BA000045_44 440001 451000
BA000045_45 450001 461000
BA000045_46 460001 4659019
Continuation (37 of 47) of BA000045 from base 360001 (BA000045 Gloeobacter violaceus P
Alignment Scores:
Pred. No.: 3,85e-17 Length: 110000
Score: 449.00 Matches: 197
Percent Similarity: 40.4% Conservative: 114
Best Local Similarity: 25.6% Mismatches: 287
Query Match: 11.1% Indels: 172
DB: 1 Gaps: 31
US-10-600-070B-2 (1-801) x BA000045_36 (1-110000)
QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
Db 89225 TTTTTCGGCGCCATGATTCGTGCTCATCCCTCTCCCTG----- 89263
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100

```


Db 89264 -----CCGCTGAGCTACTACAGATCTCGGAGTGCCTCCCAAGTCG 89305
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIysProProGlnPheGly 120
 Db 89306 ACTTACAGAGAGGAGGAGCTGCTTTTCCGAGTGCCTTGGCCGAGGCGCAGACGCGAA 89365
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleGlnAlaIleValCysGluThrLeu 140
 Db 89366 TTTTACAGCGCGCGCTCCGCTCCGCGAGACACTGCTGCGCGAAGCTGACAGCGCTG 89425
 QY 141 SerAsnProArgSerArgArgValGluTyrAsn--GluGly-----LeuLeuAsp 155
 Db 89426 GCGGATCTCCCTCCGAGCGAGCTGTACATCGAGGCGCAACAGAGATTGCTTCTCGAC 89485
 QY 156 AspGluGlnAlaThrValIleThrAspValProTyrAspIlyValProGlyAlaLeuCys 175
 Db 89486 AGCAGCCACGCGCGCTG-----GGGCTGCTGTTT 89515
 QY 176 ValIleGlnGlnGlyGlyGluThrGluIleValIleuArgValGlyGlnAlaLeuLys 195
 Db 89516 CTGTACGAAGTGGGCGCAATACCAAGCTATCGAGCGCCAGGAGAGCTTACGCGTA 89575
 QY 196 GluArgLeuProIlySerPheIySGlnAspValValIleuValMetAlaLeuAlaPheLeu 215
 Db 89576 GCGGATCAAGCCG-----GATACCGCGCTGGTGGCTCGCCCTCGCCACAG 89620
 QY 216 AspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPhe 235
 Db 89621 GCCCTGCGCCAGAGAGCTTACCGC-----CAGGCGCAATCTCACCTGCGCCACCTCGAA 89674
 QY 236 ValGluGlnAlaLeuLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 255
 Db 89675 GTTGAAGAGCTTGTGAGATTTTGC-----GGCGCGCATTCCTGAAAGCCG--GTC 89725
 QY 256 ArgAlaGlnIleAspGluThrIleGlnGlnGlnIleThrProArgTyrValIleGlnLeuLeu 275
 Db 89726 CAGCAGAGATTGAGACTCTGTCAAGCGCTGCGCGCGCGCGAGGAGATTGTCAGCTGTG 89785
 QY 276 GlyLeuProLeuGlnAspAspTyrAlaAlaIlyArgLeuAsnGlnIleuSerGlyValArg 295
 Db 89786 GCG--GAGCGCGCCAGACCCCGCTCCCGCCAGCGCGAGAGAGCAAGGCGCTGCGCG 89842
 QY 296 AsnIleLeu-----TrpSerValGlyGlyGlyValAspSerAlaLeuValGlyGly 312
 Db 89843 GCCCTGCTTCGCGAGCGCGAGCGGAGCTGAAAGCGAGCGCAACATCA-----TCGGCG 89896
 QY 313 LeuThrArgGlu-----LysPheMetAsnGlnAlaPheLeuArgMetThrAlaAla 329
 Db 89897 CTCAGCGCGAGGAATTGTTCGAGTTTTCAGATCTGCGCTGCGCGCTGAGCGGTGCGC 89956
 QY 330 GluGlnValAspLeuPhe-----ValAlaThrProSerAsnIleProAlaGluSer 346
 Db 89957 GAGCAGCAAGAAATGTTTCGAGCGCGAGCGCGAGCGCGCGCGCTGCG-----CCGCGCGCGCAG 90010
 QY 347 PheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLysPro 366
 Db 90011 -----TACCTGCGCGCGCGCGCGCAATGCGCTGCGGATTCCTGAAGGCTCGCC 90061
 QY 367 HisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnIleGlnGlnAlaLysValMetAla 386
 Db 90062 CAGTGGGTGCG 90112
 QY 387 MetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeu 406
 Db 90113 -----GACGTCAATCTG 90124
 QY 407 GluArgGlyLeuCysAlaLeuLeuIleGlyLysValAspGluCysArgMetTrpLeuGly 426
 Db 90125 GAACTGCGCGGTGCGCGCTGCTGTGTGGCCAGGTCAAGAGCCGCAAGAAACATCGAG 90184
 QY 427 LeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPheValIleGluAsn 446
 Db 90185 CGCTGCGCGAGAGAGCG-----GCGGTCAATTACATC---AAAAAC 90223

QY 447 SerAsnArgAspAspAspAspLeuProGlyLeuCysIlyLeuLeuGlnIleThrTrpLeu 466
 Db 90224 CTGTCCAGATTCGCCCGAGCTGCTGCGGGCTTGGCTCGCATACCACTGTGCTG 90283
 QY 467 AlaGlyValAlaPheProArgPheAspThrIlyAspLysIlyPheLysLeuGlyAsp 486
 Db 90284 GCGGAGAGGCTCTTTCAGGTTTTCGAGCCGCGCGCTCGGCGCACTATACCTGAGGCT 90343
 QY 487 TyrTrpAspAspProMetValIleuSerTyrLeuGlnArgValGluValValGlnGlySer 506
 Db 90344 TACTTGGCCCATCCAGAGTCCGAGCTTCCTCGAT-----GAC 90382
 QY 507 ProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGluHisValLysAlaSerAla 526
 Db 90383 CCGCAACCGCGCGCGCTCTGCGCGGAAACCGCGAGCGCGCGCGGATGCGCGCGC 90442
 QY 527 MetGlnAlaLeuGlnLysValPheProSerArgTyrThrAspArgAsnSerAlaGluPro 546
 Db 90443 GCGTCGTG-----GSCACCATCGAACC 90466
 QY 547 LysAspValGlnGluThrValPheSerValAspProValGlyAsnAsnValGlyArgAsp 566
 Db 90467 GCGCGCTG-----CTGACCGCGGTGGG----- 90490
 QY 567 GlyGluProGlyValPheIleAlaGlnAlaValArgProSerGluAsnPheGluThrAsn 586
 Db 90491 -----GCCAGGCG 90511
 QY 587 AspTyrAlaIleArgAlaGlyValSerGluSerSerValAspGluThrThrValGluMet 606
 Db 90511 ----- 90511
 QY 607 SerValAlaAspMetLeuLysGlnAlaSerValLysIleLeuAlaAlaGlyValAlaIle 626
 Db 90512 -----GCCAATCTGCTTCGCCCGAGCGCTGCTGCGGTGCGCTGCGGTGCGCAATT 90565
 QY 627 GlyLeuIleSer-----LeuPheSerGlnLysTyrPheLeuLysSerSerSer 643
 Db 90566 TTCCTCTCGCGAGCGCGCTGCTGCTACCCAGCGCGAGAGCGCGCGAGCGCTGCA 90625
 QY 644 PheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIleGlySerVal 663
 Db 90626 GCGGCGAGACC-----GTGCTCTCGGATGCGAGCTTCGCTCAATCGCGCGCGC 90679
 QY 664 ArgAlaAspAspSerGluAlaLeuProArg----- 673
 Db 90680 CCGGCGCGCGCGCGCGCGCGCTGCCACCGAGCGCTGCGCGCGAGCAACAGCTTCC 90739
 QY 674 MetAspAlaArgThrAlaGluMetIleValSerLysTyrGlnLysIleLysSerLeuAla 693
 Db 90740 ACCGAGCGCGCAATTCGCCCGC--ATGCTCAAAACTGCGCAACGCGCAAGCGAGCG 90796
 QY 694 PheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuLys 713
 Db 90797 CTGCTTCGCGAGCG 90856
 QY 714 IleTrpThrAspArgAlaAlaGluThrAlaGlnIleGlnIleuValTyrAspTyrThrLeu 733
 Db 90857 GTCGCGCAACAAAGTTCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 90916
 QY 734 LeuLysLeuSerValAspSerValThrValSerAlaAsp-----GlyThrArgAlaLeu 751
 Db 90917 AAGATCTCAAAATTCAGAGAGT-----GCCAGCGCGCGCGCGCGCGCGCGCGCG 90967
 QY 752 ValGluAlaThrIleGlnGlnLysSerAlaCysLeu-----SerAspLeuValHisProGlu 769
 Db 90968 GTGACGCGCGAGCTACCGAGTGTGCGCATCTTACACCGAGCAATAGCTGCGCGCTG 91027
 QY 770 AsnAsnAlaThrAspValArgThrTyrThrArgTyrGluValPheTrpSerLysSer 789
 Db 91028 CGCTCTACGAC-----GCGCCCTACCGGAGCGGTATCTCCCTGTCAAGAGCGCTGCC 91081

Or	790	GI:1791751	ethThnGluGlySerValIleu	799
Db	91082	GGCTGGCGGATCGAGAGATGAAAGTGTTG	91111	
RESULT 15				
AF421196				
LOCUS				
DEFINITION	AF421196	2469 bp	DNA	linear BCT 10-SEP-2002
	Synechococcus sp. PCC 7942	cell division protein Ftn2 gene,		
	complete cds.			
ACCESSION	AF421196			
VERSION	AF421196.1	GI:16226083		
KEYWORDS				
SOURCE				
ORGANISM				
	Synechococcus elongatus PCC 7942			
	Synechococcus elongatus PCC 7942			
	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.			
REFERENCE	1 (bases 1 to 2469)			
AUTHORS	Koksharova,O.A. and Wolk,C.P.			
TITLE	A Novel Gene That Bears a DnaJ Motif Influences Cyanobacterial Cell Division			
JOURNAL	J. Bacteriol.	184 (19),	5524-5528	(2002)
PUBMED	12218043			
REFERENCE	2 (bases 1 to 2469)			
AUTHORS	Koksharova,O.A. and Wolk,C.P.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-SEP-2001)	Plant Research Laboratory, Michigan State University, DOE Plant Research Laboratory, Bact Lansing, MI 48824,		
	USA.			

RES source	Location/Qualifiers
1. 2469	/organism="Synecococcus elongatus PCC 7942" /mol_type="genomic DNA" /strain="PCC 7942" /db_xref="taxon:1140" 319. .2214
CDS	

```

/codon_start=1
/translate_table=11
/product="cell division protein Fcn2"
/protein_id="P41167.1"
/db_xref="GI:16226084"
/translation="MRIPLDYRILCVGVQASADKLAEVYDRINQSPSHFSELALQ
ARQLLEAIAELSDPEODRYDRIRFPFGGLEALEPSELEDMORI GALLILLEGEEY
DRVQALAEELL PDVQASAEVRDOPFARGI IALALSQGLRECRQGLYQAOAHQGL
RQGLALADHQRPPELSRTILNBOGGLRYYRILERLAQVLTZDSDRQGLLLQAMLDD
RQGLEPGDDSGSLTDNFMFLQRIKYLTLAEQQLPFESARSPSPASFFRCYITL
IARGCDHPSLIHRAISLLHELEKSRMDVHLEQALISLLAQPEAEALVQSDDEETL
LSQRLAALGALIVGLCRFETETMLATVPFDPIKRTIAPLQYPEDPDVQYTLAPL
IVELPSDMLPTPLPVEPLVRSLSLAKELPTPAAGVAVPRRRRSDSEKAPATKARK
LPLPRIGLVGVVVLGGGTCWAMKRSRSPTPTPQVTLPAVPAASPAATVALLD
RAQAEIVLQNLWLAKAALGAPQYDRDRLATVLTGEVLTQWGFSSQANTQLTQSPDHH
KLTVDSVQLSDGDQRAVQAKVDEVEQYRBDQLLETRDRDGLVIYRQLVRENNIMKLI
ASISILVR"

```

ORIGIN	
Alignment Scores:	
Pred. No.:	5 4e-19
Score:	446.00
Percent Similarity:	38.8%
Best Local Similarity:	23.7%
Query Match:	11.0%
DB:	1
Gaps:	
Length:	2469
Matches:	203
Conservative:	129
Mismatches:	252
Indels:	232
Gaps:	33

US-10-600-070B-2 (1-801) x AF421196 (1-2469)

OY 4 LeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCysArgLeuProProAlaThr 23
 ||||| :||| :||| :||| :||| :||| :||| :|||
Db 161 CTCTCAAGATATGGGTTTTCGTCCGCG--TATTCCCATCATCCCCCTGTCTGCT 217
 ||||| :||| :||| :||| :||| :||| :||| :|||
OY 24 ThrIleValLeuArgSerSerHisAsnThrSerThrIleCysSerAlaSerLysTrpAla 43
 ||||| :||| :||| :||| :||| :||| :||| :|||
Db 218 ACGGAGATTAAC-GCCGATCACTCAACAAGA--TTGTAAGTTGACGGTCAAGTGCGGAT 273
 ||||| :||| :||| :||| :||| :||| :||| :|||
OY 44 AspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSerPheAlaThr 63
 ||||| :||| :||| :||| :||| :||| :||| :|||

Db 274 GATGAAGTCGGCTCAAGCTGGCGGATCCGGATCTGTGGTGGTCT----- 318
 Oy 64 AlaIhThrThraIaThrLeuValSerLeuProProSerIleAspArgProGluArgHis 83
 Db 318 ----- 318
 Oy 84 ValProIleProIleAspPheTyGlnValLeuGlyValaGlnThrHisPheLeuThrAsp 103
 Db 319 GTGGCTATTCTCTCGATTACTACCGAAATCTCTGTGTGGCGCGCAAGCTCGGCAGAC 378
 Oy 104 GilyleArgArgAlaPheGluValaArgValSerIleAspProGlnIlePheSerAsp 123
 Db 379 AAACCTGCCAAGAGCTACCGCGATCGGCTCAACCAATCGCCCTCCATGATTTTCAGAG 438
 Oy 124 AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCyseGluThrIleSerAspPro 143
 Db 439 CTGGCATTTGACGGCGCGCGCACTCTCGAAGCACAGATTGCTAGCTGAGGAATCCC 498
 Oy 144 ArgSerArgArgGluTyArgAnglu-----GlyLeuLeuAspArgGlu 158
 Db 499 GAACAGCGCGATCGCTACGATCGCGGCTTTTTCAGGGCGGCTCTGGAAGCATTGAACCA 558
 Oy 159 AlaIhThrValIleThraAspValProTyrAspIleValProGlyAlaLeuCyseValLeuGln 178
 Db 559 AGCCTGAAGACTCGAAGAC-----TGGCAGGAAATT---GGAGCCCTGCGATCTGTG 609
 Oy 179 GluGlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeu----- 194
 Db 610 GAATTGGGGGAATACGATCGGCTTCCCACTGGCTCAAGAACTCTCCGACAGTACAGAC 669
 Oy 195 -----LysGluArgLeuProLysSerPheIleGlnAspValValLeuValMetAlaLeu 212
 Db 670 GCGAGCCGCAAGATGACGCGATCACTTGGCGGGGTATATCGCTTGGCGATGCGACTA 729
 Oy 213 AlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGly 232
 Db 730 TCCAGCAATCCCTCGGTGAGATGCGCTGAG----- 762
 Oy 233 TyrGluPheValGluGluAlaLeuLysLeuLeuGlnGluGlyAla----- 248
 Db 763 -----CAGGGTCTGTACCAACAGCGCGCCGACAGACTTTGAC 798
 Oy 249 -----SerSerLeuAla-----ProAspLeuArgAlaGlnIleAsp 260
 Db 799 CGCAGCCAGTTCGCCCTAGCCGATCATCAGCGCTTCTTGAACTGAAGTCGAACCTGCAC 858
 Oy 261 GluThrLeuGluGluIleThrProArgTyValLeuGlnLeuLeuLeuGlyLeuProLeuGly 280
 Db 859 CAAGAACAAGACACAGCTACGCGCTTATCGCATTTTGGAGCGGTTGGCCAGCCCTTAGCT 918
 Oy 281 AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeu----- 298
 Db 919 GCCGAT-----AGCGATCGGCACAGAGGGTTTCTGTGTGTCAGCGAGATGTTGAGAGAC 972
 Oy 299 ---TrpSerValGlyGlyGlyValAserAlaLeuValGlyGlyLeuThrArgGluLys 317
 Db 973 CGCAGCGGCATTGAAGCGCTCGGGAGATGAGC-----TCGGGGCTGACCTTGATTAAC 1026
 Oy 318 PheMetArgnGluAlaPheLeuArg-----MetTyrAlaGlnGlnVal 332
 Db 1027 TTTTGG-----ATGTTTCTCCAGCAAAATTCGCGCGATCTGACCCCTGGCGAAGACAG 1080
 Oy 333 AspLeuPhe-----ValAlaThrProSerAsnIleProAlaGluSerPheGluVal 349
 Db 1081 TTGCTGTTTGAATGGGAAGCGCGTGGCGCTCG-----CGGCGTGGAGGCTT----- 1128
 Oy 350 TyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLysIleProHisLeuLeu 369
 Db 1129 ---TTTGGCTGCTACACCTGATGTCGGGGGGCTTTTGGCATTCACAACTCTGATATC 1185
 Oy 370 GlnAspAlaAspLysGlnPheGlnGlnLeuGlnAlaLysValMetAlaMetGluIle 389
 Db 1186 CATGCGCGCAGCTTGCTCTGCGATGAATCAAGAC----- 1221

```

Qy 390 ProAlaMetLeuThrArgAsnTrpGluLeuPheGlyLeuGluArgGly 409
Db 1222 -----CGCATGATGCGCATCGAACAGCGC 1248
Qy 410 LeuCyAlaLeuLeuLeuLeuValAspGluCysArgMetTrpLeuGlyLeuAspSer 429
Db 1249 ATCGCAGAGCTATTGCTCGGACAGCCCGAAGAGCTAGGGCGCTACTC---GTCACAGC 1305
Qy 430 GluAsp-----SerGlnTrpArgAsnProAlaIleValGluPheValLeuGlu 445
Db 1306 CAAGATAGAGAAACCTCAGCCCAATCCGTGCTTACCCCAAGGAGGAGCCCTG- 1359
Qy 446 AsnSerAsnArgAspAsnAspAspLeuProGlyLeuCyAlaLeuLeuGluTrp 465
Db 1360 -----ATCGTCGATTTGCGCATTCAGCAAGAACTGG 1392
Qy 466 LeuAlaGlyValAlaPheProArgPheArgAspTrpLysAspLysPheLysLeuGly 485
Db 1393 CTAGCGACCAAGGATTTCCGGAATTCGCGACTCAAGAAAGACTGGCGCGCTGCAG 1452
Qy 486 AspTyrTrpAspAspProMetValLeuSerTyrLeuGluArgValGluValGlnGly 505
Db 1453 CCTACTTTCAGCAGCCCGATGTCAGACCTATCTGATGCGATCGTGAATTG----- 1506
Qy 506 SerProLeuAlaAlaAlaIleThrMetAlaArgIleGlyAlaGluHisValLysAlaSer 525
Db 1506 ----- 1506
Qy 526 AlaMetGluAlaLeuGlnLysValPheProSerArgTyrThrAspArgAsnSerAlaGlu 545
Db 1506 ----- 1506
Qy 546 ProLysAspValGlnGlnTrpValPheSerValAspPro----- 558
Db 1507 CCTCGCATTTGATGCGCAAGCCGCTACCCGTTGAGCCGCTGAGTGCATGCTGCTG 1566
Qy 559 ValGlyAsnAsnValGlyArgAspGlyGluProGlyValaPheIleAlaGluAlaValArg 578
Db 1567 CTGGCCAAAGAACTGCGCAGCCCAAGCAAGCCCTGCT-----GTAGCT 1608
Qy 579 ProSerGluAsnAspGlnTrpAsnAspTyrAlaIleArgAlaGlyValSerGluSerSer 598
Db 1609 CCAACCCCTCGCGCGCGCTCGCGCATCGCTCCGAACGCTCTGCTGCGCAGGCCAAAGC 1668
Qy 599 ValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLys 618
Db 1669 TTGCCCCCTGCGATTGCTTGGGGGTT-----GTGGTG 1704
Qy 619 IleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeu 638
Db 1705 GTTCTCGCGCGTGAACAGGGGTT-----TGCGCTTGG 1737
Qy 639 LysSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAla 658
Db 1738 CGATCGCGTTCCAAT-----TTCACCCGCGCG 1764
Qy 659 ThrIleGlySerValArgAlaAspAspSerGluAlaLeuPro----- 672
Db 1765 ACCCCGCCCCCGTGGTTCAAAAGCTGCTGAGGCGGATCTGCCCTTGCAGCCGCGCA 1824
Qy 673 -----ArgMetAspAlaArgThrAlaGluAsnIleValSerLysTrpGlnLysIle 689
Db 1825 GTTACCGTTCCTCATCGGCTCAGGCTGAAGAACTGTGTGCAAAACTGTGGCGCGCT 1884
Qy 690 LysSerLeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGly 709
Db 1885 AAAGCTGCAAGCTTGGGCTCTAATACATCGGATCGCTTACGACGAGTGAACCGGT 1944
Qy 710 ArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAla-----GlnLeuGlyLeu 727
Db 1945 GAGGTTCTGCAACTTGGCAGGGGTTTCTTACGACAGGCAAGCAACCGAGTCAATCA 2004

```

```

Qy 728 ValTyrAspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGly 747
Db 2005 CAGTTCGATCAC-----AAGTTAACCGCTGACTCACTTCACTCACTGACGGTGAT 2055
Qy 748 ThrArgAlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHis 767
Db 2056 CAACGAGCAGATGTCMAAGCCAAAGCTGATGAA-----GTTGAGCAGGTCTAT 2103
Qy 768 ProGluAsnAsnAlaThrAspValArgThr-----TyrThrThrArgTyrGluVal 784
Db 2104 CGAGCGCAGCAAGCTGCTCGAAACGCGCGAGATTGGGCTTGGTGAATCCCTACAGCTC 2163
Qy 785 PheTrpSerLysSerGlyTrpLysIleThrGluGlySerValLeu 799
Db 2164 GTGCGCAGAAACATCTGAAATTTGCTTCGATTTAGTTGGTG 2208

```

Search completed: February 21, 2006, 18:40:01
Job time : 11481 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 13:54:21 : Search time 11793 Seconds
(without alignments)
11597.163 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagcctctgacgtcagcgt.....gctctgtctctgacataca 2406

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_dr: *
9: gb_to: *
10: gb_sts: *
11: gb_sy: *
12: gb_un: *
13: gb_vi: *
14: gb_hlg: *
15: gb_pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2402.8	99.9	2438	AY221469	AY221469 Arabidops
2	2399.6	99.7	2436	AY150450	AY150450 Arabidops
3	2399.6	99.7	2637	AY091075	AY091075 Arabidops
4	1724	71.7	3668	AY221468	AY221468 Arabidops
5	1720.8	71.5	3664	AY221467	AY221467 Arabidops
6	1717.6	71.4	85791	AB016888	AB016888 Arabidops
7	243.6	10.1	133779	AC158210	AC158210 Medicago
8	180.8	7.5	110000	AP008208_011	Continuation (12 o
9	180.8	7.5	150462	AP004885	Oriza sat
10	165.4	6.9	550	DQ022571	Beta vulg
11	163.4	6.8	184752	AC155597	AC155597 Zea mays
12	144.2	6.0	117818	AC160013	AC160013 Medicago
13	62	2.6	221756	AC094893	AC094893 Rattus no
14	62	2.6	232974	AC114193	AC114193 Rattus no
15	60.8	2.5	200412	AC115666	AC115666 Rattus no
16	60.8	2.5	253094	AC113774	AC113774 Rattus no
17	60.2	2.5	232373	AC134805	AC134805 Rattus no
18	59.6	2.5	193341	AC134500	AC134500 Rattus no

19	59.2	2.5	96573	14	AC136672	AC136672 Rattus no
20	59.2	2.5	175351	14	AC117900	AC117900 Rattus no
21	59.2	2.5	233591	14	AC121387	AC121387 Rattus no
22	59.2	2.5	243553	14	AC096320	AC096320 Rattus no
23	59.2	2.5	247341	14	AC099178	AC099178 Rattus no
24	59.2	2.5	312143	14	AC093509	AC093509 Rattus no
25	59.2	2.5	346208	14	AC123509	AC123509 Rattus no
26	57.6	2.4	7218	6	166494	166494 Sequence 14
27	57.6	2.4	255678	14	AC099432	AC099432 Rattus no
28	57.6	2.4	313096	14	AC131879	AC131879 Rattus no
29	57.2	2.4	98521	9	AC094507	AC094507 Rattus no
30	57	2.4	263036	14	AC105597	AC105597 Rattus no
31	56.8	2.4	110000	14	CR954204_0	CR954204 Oetococ
32	56.4	2.3	282355	14	AC104401	AC104401 Rattus no
33	56	2.3	143976	14	AC141543	AC141543 Rattus no
34	56	2.3	246199	14	AC119558	AC119558 Rattus no
35	56	2.3	252996	14	AC130153	AC130153 Rattus no
36	56	2.3	253772	14	AC109777	AC109777 Rattus no
37	55.6	2.3	242516	14	AC106523	AC106523 Rattus no
38	54.6	2.3	165849	14	AC114251	AC114251 Rattus no
39	54.6	2.3	231716	9	AC135409	AC135409 Rattus no
40	54.6	2.3	234812	14	AC111482	AC111482 Rattus no
41	54	2.2	110000	14	AC129796_0	AC129796 Rattus no
42	54	2.2	240800	14	AC095964	AC095964 Rattus no
43	54	2.2	274697	14	AC098240	AC098240 Rattus no
44	53.8	2.2	205796	14	AT713891	AT713891 Mus muscu
45	53.8	2.2	215524	9	AL603913	AL603913 Mouse DNA

ALIGNMENTS

RESULT 1	AY221469	2438 bp	mRNA	linear	PLN 05-AUG-2003
LOCUS	AY221469				
DEFINITION	Arabidopsis thaliana division protein (ARC6) mRNA, complete cds;				
ACCESSION	AY221469				
VERSION	AY221469.1	GI:33436353			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons; rosids; eustosida II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1. (bases 1 to 2438)				
AUTHORS	Viltha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Ert,H. and Oesteryoung,K.W.				
TITLE	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary				
JOURNAL	Descendant of the Cyanobacterial Cell Division Protein Ftn2				
PUBMED	12897262				
REFERENCE	2. (bases 1 to 2438)				
AUTHORS	Viltha,S., Koksharova,O., van Ert,H., Froehlich,J.E. and Oesteryoung,K.W.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA				
FEATURES	Location/Qualifiers				
source	1. .2438				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/db_xref="taxon:3702"				
	/ecotype="Massilewskija"				
	1. .2438				
	/gene="ARC6"				
	/note="synonym: At5g42480"				
	2. .2407				
	/gene="ARC6"				
	/codon_start=1				
	/product="division protein"				
	/protein_id="AA018646.1"				
	/db_xref="GI:33436354"				
CDS					

```
/cranelation="MEALSHVIGLSIPOLCRPPATTKLRSHNTSTTICSAKMWAD
RLSDNPFSSSSSPATTTATVPSIDRERHPPIIDRYOVLQADTHELT
DGRBAPAPVPOGSPDDALISROLOAACTLSNPSRBRBYNGLDDEGAT
VITDVPMKVPKALCYLOEGREIYLRGEBALIKERLPSFQDVIYUWALAFLDS
RAMALDPDPITIGYEFVEBALILOEBGASLAPDLRAQIDTLEITPRYVLELIG
LPLGDDYAAKRLNGSLGVRNLIWSVGGASALVGLTRKFNEMEPFLRTAAEQVL
FVTPENIIPAESFEVVALALVAQAFIGKPHILODADKFOLOQAKMAEITPAM
LYOTRNNBIIDPFLERGLCALIGKVDCEMMWGLDSEDSOYRPAIIVERFVLENSRD
DUNDLPGLCTLEETWGLGAVPRPRDCKKFLGPDYDDPMVLSTLEREVVQGSPL
AAATATVRIAEHVYKASAMOLOKVPFSPRTIDRNSAPRHOVQETVSVDPGNVNRD
GEBGVFIAEVRSENFTNDYAIRAGVSSSVDETVEMSVAMELKESVKILAGV
AIGLISLFSQKYLKSSSFORKDMWSMESDVATIGSVADQSEALPRMDATTAANI
VSKMOKISLAFDPDRHIEMLPEVLDGRMLKWTDRBAAEATQOLGLVYDTLLSTVDS
VWVSADGTRALVETATLEESACLSLDIVHPENNATDVRYTTRYEVFWSKSMKITGESV
LAS"
misc_difference 474
/gene="ARC6"
/notes="results in Glu to Glu substitution"
/replace="a"
misc_difference 1831
/gene="ARC6"
/notes="results in Glu to Asp substitution"
/replace="t"
ORIGIN
Query Match 99.9%; Score 2402.8; DB 15; Length 2438;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAAAGCTCTGATCAGTCGTCGATGTCCTCTCCCATTCCAATTATGCGCATTAACA 60
Db 2 ATGAAAGCTCTGATCAGTCGTCGATGTCCTCTCCCATTCCAATTATGCGCATTAACA 61
QY 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTTACACTATCTGTCGCGCACG 120
Db 62 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTTACACTATCTGTCGCGCACG 121
QY 121 AAATGGGCGGACGCTCTCTCTCGACTTCAATTTCACCTCCGATCTCTCTCTCTCC 180
Db 122 AAATGGGCGGACGCTCTCTCTCGACTTCAATTTCACCTCCGATCTCTCTCTCTCTCC 181
QY 181 TTGCGCACCGCCACCAACGCGCACTCTCTCTCTGCGCACCATCTATTTGATCGTCCC 240
Db 182 TTGCGCACCGCCACCAACGCGCACTCTCTCTCTGCGCACCATCTATTTGATCGTCCC 241
QY 241 GAAAGCCACGTCCTCCATCCCATTTGATTTCTTACCAAGTATTGAAGCTCAAAACATTTT 300
Db 242 GAAAGCCACGTCCTCCATCCCATTTGATTTCTTACCAAGTATTGAAGCTCAAAACATTTT 301
QY 301 TTAAACGATGAATCAAGAGAGCATTTGAAGCTAGGGTTTGAACGCGCGCAATTCGGT 360
Db 302 TTAAACGATGAATCAAGAGAGCATTTGAAGCTAGGGTTTGAACGCGCGCAATTCGGT 361
QY 361 TTGACGACGACGCTTTAATCAGCGGACGAGATTCCTTCAAGCTGTCTTGGAATCTTG 420
Db 362 TTGACGACGACGCTTTAATCAGCGGACGAGATTCCTTCAAGCTGTCTTGGAATCTTG 421
QY 421 TCTAATCTCGGCTTGAAGAGAGTACATGAGTCTTCTTGATGATGAAGAGCTTACA 480
Db 422 TCTAATCTCGGCTTGAAGAGAGTACATGAGTCTTCTTGATGATGAAGAGCTTACA 481
QY 481 GTCTATCTGATGTTCTCTGGGATTAAGGTTCTGGGGCTCTCTGTGATGATGAAGAGGT 540
Db 482 GTCTATCTGATGTTCTCTGGGATTAAGGTTCTGGGGCTCTCTGTGATGATGAAGAGGT 541
QY 541 GGTGAGACTGAGATGATCTTCGGGTTGGTAGAGCTCTGCTTAAGAGAGAGTTGCTTAAG 600
Db 542 GGTGAGACTGAGATGATCTTCGGGTTGGTAGAGCTCTGCTTAAGAGAGAGTTGCTTAAG 601
QY 601 TCGTTAAGCAAGATGATGTTAAGTATGAGCGCTTGCTTGATGATGATGATGATGATGAT 660
Db 602 TCGTTAAGCAAGATGATGTTAAGTATGAGCGCTTGCTTGATGATGATGATGATGATGAT 661
```

```
QY 661 GCTATGAGCATTTGATCAGCTGATTTTATTTACTGCTTATGATGATTTGTGAGAAAGCTTTG 720
Db 662 GCTATGAGCATTTGATCAGCTGATTTTATTTACTGCTTATGATGATTTGTGAGAAAGCTTTG 721
QY 721 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 722 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
QY 781 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGTGAGACTTATGAGCTTATGAGCTTATG 840
Db 782 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGTGAGACTTATGAGCTTATGAGCTTATG 841
QY 841 GATGATTAAGCTGCGAAGAGACTTAAATGCTTAAAGCGGTGTGCGGAATATTTGTGTCT 900
Db 842 GATGATTAAGCTGCGAAGAGACTTAAATGCTTAAAGCGGTGTGCGGAATATTTGTGTCT 901
QY 901 GTTGAGAGAGAGTGAACATCAAGCTCTGTTGGGGGTTTACCCGCTGAGAAATTATGAT 960
Db 902 GTTGAGAGAGAGTGAACATCAAGCTCTGTTGGGGGTTTACCCGCTGAGAAATTATGAT 961
QY 961 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGAGTGAATCTTTTGTAGCTACCCGAGC 1020
Db 962 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGAGTGAATCTTTTGTAGCTACCCGAGC 1021
QY 1021 AATATTTCCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACACTTCTTGTGCTCAAGCT 1080
Db 1022 AATATTTCCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACACTTCTTGTGCTCAAGCT 1081
QY 1081 TTTATTTGATGAAGAGCAGACCTTTTACAGAGAGTGAATGAAGAAATTCGCAACTTCAG 1140
Db 1082 TTTATTTGATGAAGAGCAGACCTTTTACAGAGAGTGAATGAAGAAATTCGCAACTTCAG 1141
QY 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGGAATTAATG 1200
Db 1142 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGGAATTAATG 1201
QY 1201 GAGATGAGCTTCGCTTACGAAGAGGAGCTGTGACCTGCTTATGAGCAAAAGTTGATGA 1260
Db 1202 GAGATGAGCTTCGCTTACGAAGAGGAGCTGTGACCTGCTTATGAGCAAAAGTTGATGA 1261
QY 1261 TGGCGTATGCTGTTGGGCTTACAGAGAGAGATTCACAAATATGAGAAATTCAGATATG 1320
Db 1262 TGGCGTATGCTGTTGGGCTTACAGAGAGAGATTCACAAATATGAGAAATTCAGATATG 1321
QY 1321 GAGTTGTTTGAAGAAATTCAAATCGATGAGACATGATGATCTCCCTGAGCTATGCAAA 1380
Db 1322 GAGTTGTTTGAAGAAATTCAAATCGATGAGACATGATGATCTCCCTGAGCTATGCAAA 1381
QY 1381 TTGTTGAAAACCTGTTGGAGAGGTTGCTTCTCTAGTTTCAAGACACCAAAAGATTA 1440
Db 1382 TTGTTGAAAACCTGTTGGAGAGGTTGCTTCTCTAGTTTCAAGACACCAAAAGATTA 1441
QY 1441 AAATTTAACTCGGGGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1442 AAATTTAACTCGGGGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 1501
QY 1501 GAGTATGTTCAAGGTTCTCTTTAGCTGCTGCTGCACTTATGAGCAAGATTTGAGAGCGGAG 1560
Db 1502 GAGTATGTTCAAGGTTCTCTTTAGCTGCTGCTGCACTTATGAGCAAGATTTGAGAGCGGAG 1561
QY 1561 CATGTGAAGCTGATGCTATGACAGGCACTGCAAGAAATTTTCTTCCCTGATATACAGAT 1620
Db 1562 CATGTGAAGCTGATGCTATGACAGGCACTGCAAGAAATTTTCTTCCCTGATATACAGAT 1621
QY 1621 AGAACTCGGCTGAACCCAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1622 AGAACTCGGCTGAACCCAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
QY 1681 AACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1682 AACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
QY 1741 GAAAACTTTGAAAATTAATGATTAATGCAATTCGAGCTGGGGCTTCAGAGAGATGAGCTTTGAT 1800
```

```

Db      1742  GAAAACTTTGAACTATGATATGCAATTCGAGCTGGGGTCTCAGAGAGTGGTTGAT 1801
Qy      1801  GAAACTACTCTTGAAGATGTCCTGCTGATATGTTAAAGAGCAATGCAAGATCTCTA 1860
Db      1802  GAAACTACTCTTGAAGATGTCCTGCTGATATGTTAAAGAGCAATGCAAGATCTCTA 1861
Qy      1861  GCTGCTGCTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1920
Db      1862  GCTGCTGCTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1921
Qy      1921  AGCTCATCTTTTCAACGCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCA 1980
Db      1922  AGCTCATCTTTTCAACGCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCA 1981
Qy      1981  GGGTCAGTCAAGAGTCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCAAT 2040
Db      1982  GGGTCAGTCAAGAGTCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCAAT 2041
Qy      2041  AATATGATATCAAGAGTCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCA 2100
Db      2042  AATATGATATCAAGAGTCAAGATATGCTTCTTATGCAATGCAATGCAATGCAATGCA 2101
Qy      2101  GAAATGTTACCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db      2102  GAAATGTTACCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
Qy      2161  GAAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db      2162  GAAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2221
Qy      2221  GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db      2222  GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2281
Qy      2281  TGTCTATCTGATTTGGTTCATCCAGAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db      2282  TGTCTATCTGATTTGGTTCATCCAGAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341
Qy      2341  AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db      2342  AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2401
Qy      2401  TCATAA 2406
Db      2402  TCATAA 2407

```

```

RESULT 2
AY150490 2436 bp mRNA linear PLN 23-SEP-2002
LOCUS Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete
DEFINITION
cds.
ACCESSION AY150490
VERSION AY150490.1 GI:23297711
KEYWORDS FLU CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukarjaya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2436)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Kimura,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,M.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R., and Theologis,A.
Arabidopsis Open Reading Frame (ORF) clones
REFERENCE
2 (bases 1 to 2436)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,

```

```

TITLE
JOURNAL
COMMENT
The Salk, Stanford, PEGC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,
Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shim,M., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R., and Theologis,A.
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Natsuka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
FEATURES
source
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to Genbank.
Location/Qualifiers
1..2436
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U13935"
/ecotype="Columbia"
/note="This clone is in PUNI 51."
1..2436
/gene="At5g42480"
1..2406
/gene="At5g42480"
/evidence="experimental"
/product="unknown protein"
/protein_id="PANI2907.1"
/db_xref="GI:23297712"
/translation="MEALSHVIGLSPPOLCRPPTATKLRSHNTSTTICGASKAD
RLSDPNFTSDSSSFATTTATVATVPPSPIDRBRVPIIDFYOVYGAOTHPLT
DGIIRAFARVSKPPGSGSDALISROILCAACTLSNPSRRRNGILDDDEAT
VITDVPMDKVPGLCVLQEGTEIVLRVGBALKEKLRPSKQDYVVALAFDVS
RAMALDPDPDFTGVEFEVBEALQLQEGASLAPLRLQIDTLEBILPRVYLELG
LPAGDYPAKRLNGLSGVNRIILMSVGGASALVGLTRKFNBAFLMTAAEOVDL
FVATPSNIAAESFEVYVALVAQAFICGKPHLDADKQFOLOAKYMAEIPAM
LYDTRNWEIDRGLERGLCALLIGKDEKRMWIGLSDSOYRNPILIVERVENSND
DNDDEGLCKLETLVAGVYFFRPTKRCGLGQYDDPWLSTLEREVYQSPRL
AAAAAARLGAELVKASAMQALQKVPKSTYDKNASBPVOVSTVSPVGVNGVD
GEGVYIAVRSSENFENDYAIRAGVSSSVDETIVMSVADMLKESVKTLAGV
ATGLSLFQKTFKSSSFORKDWSMESDVATIGSVRADDSBALPRMDATANI
VSKWQKISLAFQPDHRIEMLPEVDGRMLKWTDRASFAQGLVADYVTLKTSVDS
VYVSAQGTALVATLEESACSLDVHPENNMTDVRTYTRREVFVSKSGWKITBSY
LAS"
2407..2436
/gene="At5g42480"
2433..2434
/gene="At5g42480"
/note="compared to genomic sequence"
/replace="ctg"
ORIGIN
Query Match 99.7%; Score 2399.6; DB 15; Length 2436;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGGAAGCTCTGATGACGCGGCAATTGCTCTCCCAATTGCAATTATGCGGATTACCA 60
Db 1 ATGGAAGCTCTGATGACGCGGCAATTGCTCTCCCAATTGCAATTATGCGGATTACCA 60
QY 61 CCGGACGACGAAAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCTCCGACG 120
Db 61 CCGGACGACGAAAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCTCCGACG 120
QY 121 AAATGGGCGGACCGCTCTTCTCTCCGACTTCAATTTCACCTTCGCAATTCCTCTCC 180
Db 121 AAATGGGCGGACCGCTCTTCTCTCCGACTTCAATTTCACCTTCGCAATTCCTCTCC 180
QY 181 TTGCGCACGCGCCACACACCGGCACTCTGCTCTGCGACCACTATATGATGCTCC 240
Db 181 TTGCGCACGCGCCACACACCGGCACTCTGCTCTGCGACCACTATATGATGCTCC 240
QY 241 GAAGCGACGCTCCCAATCCCAATTGATTTCTACCAAGTATTAGAAGCTCAAAACAATTC 300
Db 241 GAAGCGACGCTCCCAATCCCAATTGATTTCTACCAAGTATTAGAAGCTCAAAACAATTC 300
QY 301 TTAAACGATGGAATCAGAAAGCAATTCGAAAGCTAGGCTTCCAAACCGCCGCAATTCG 360
Db 301 TTAAACGATGGAATCAGAAAGCAATTCGAAAGCTAGGCTTCCAAACCGCCGCAATTCG 360
QY 361 TTACAGGACGACGCTTTAACTCAGCCGCGACAGATCTTCAAGCTGCTTGCGAAACTCTG 420
Db 361 TTACAGGACGACGCTTTAACTCAGCCGCGACAGATCTTCAAGCTGCTTGCGAAACTCTG 420
QY 422 TCTAATCTCGGTCTAGAAAGAGATCAATGAAGGCTTCTTGATGATGAAAGCTAC 480
Db 422 TCTAATCTCGGTCTAGAAAGAGATCAATGAAGGCTTCTTGATGATGAAAGCTAC 480
QY 481 GTGATCACTGATGTTCTTGGAATAGGTTCTGCGGCTCTCTGTGATTTGCAAGAGT 540
Db 481 GTGATCACTGATGTTCTTGGAATAGGTTCTGCGGCTCTCTGTGATTTGCAAGAGT 540
QY 541 GGTGAGACTGAGATGATGTTCTGCGGTTGAGAGGCTCTGCTTAAAGAGAGGTTGCCCTAG 600
Db 541 GGTGAGACTGAGATGATGTTCTGCGGTTGAGAGGCTCTGCTTAAAGAGAGGTTGCCCTAG 600
QY 601 TCGTTTAAAGCAATGATGTTTATGATAGGCGCTTGCTGATGATCTGAGGAGT 660
Db 601 TCGTTTAAAGCAATGATGTTTATGATAGGCGCTTGCTGATGATCTGAGGAGT 660
QY 661 GCTATGCGATTTGATTCACCTGATTTTATCTGCTTATGATGATTTGTTGAGAAAGCTTTG 720
Db 661 GCTATGCGATTTGATTCACCTGATTTTATCTGCTTATGATGATTTGTTGAGAAAGCTTTG 720
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACGCTTGACACCGGATTTTACGTCGACAAATTGAT 780
Db 721 AAGCTTTTACAGAGAGAGAGCAAGTACGCTTGACACCGGATTTTACGTCGACAAATTGAT 780
QY 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGT 840
Db 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGT 840
QY 841 GATGATTTACGCTGGAAGAACTAATGTTTAAAGCGGTGCGGAATTTTGTGCTCT 900
Db 841 GATGATTTACGCTGGAAGAACTAATGTTTAAAGCGGTGCGGAATTTTGTGCTCT 900
QY 901 GTTGAAGAGAGTGAAGATCACTGCTTGTGGGGGTTTGAACCCGTGAGAAATTATGAT 960
Db 901 GTTGAAGAGAGTGAAGATCACTGCTTGTGGGGGTTTGAACCCGTGAGAAATTATGAT 960
QY 961 GAGGCGTTTTTACGAATGACGCTGAGAGGTTGATCTTTTGTAGTACCCCAAGC 1020
Db 961 GAGGCGTTTTTACGAATGACGCTGAGAGGTTGATCTTTTGTAGTACCCCAAGC 1020
QY 1021 AATATTTACAGAGAGTATTTGAAGTTTACGAAGTTTCACTTGTCTTGTGGCTCAAGCT 1080
Db 1021 AATATTTACAGAGAGTATTTGAAGTTTACGAAGTTTCACTTGTCTTGTGGCTCAAGCT 1080

QY 1081 TTTATTTGTAAGAACCAACACCTTTTACAGATGCTGATTAACCAATTTCCAGACTTCAG 1140
Db 1081 TTTATTTGTAAGAACCAACACCTTTTACAGATGCTGATTAACCAATTTCCAGACTTCAG 1140
QY 1141 CAGGCTAAGGTAAAGCTTATGAGATTCCTGGGATGTTGTATGATACAGGAATATTTG 1200
Db 1141 CAGGCTAAGGTAAAGCTTATGAGATTCCTGGGATGTTGTATGATACAGGAATATTTG 1200
QY 1201 GAGATGACTTTCGCTTATGAAAGGGGACTCTGTGCACTGCTTATAGGCAAGTTGATGA 1260
Db 1201 GAGATGACTTTCGCTTATGAAAGGGGACTCTGTGCACTGCTTATAGGCAAGTTGATGA 1260
QY 1261 TCCCGTATGCTGCTTGAACAGTGAATTCACATATAGGAATCAGCTATTTG 1320
Db 1261 TCCCGTATGCTGCTTGAACAGTGAATTCACATATAGGAATCAGCTATTTG 1320
QY 1321 GAGTTGTTTGAAGAAATTCAAATGCTGATCAATGATGATCTCCGTGACATATGCA 1380
Db 1321 GAGTTGTTTGAAGAAATTCAAATGCTGATGATCAATGATGATCTCCGTGACATATGCA 1380
QY 1381 TTGTTGAAACCTGGTTGCGAGGGGTTCTTCTCTAGGTTGAGAGACCAAAAGATTA 1440
Db 1381 TTGTTGAAACCTGGTTGCGAGGGGTTCTTCTCTAGGTTGAGAGACCAAAAGATTA 1440
QY 1441 AAATTTAACTCGGGGACTACTATGATGATCTGATGTTTGAATTTACTTGAAGAGT 1500
Db 1441 AAATTTAACTCGGGGACTACTATGATGATCTGATGTTTGAATTTACTTGAAGAGT 1500
QY 1501 GAGGTATGTTCAAGGTTCTCTTTATGCTGCTGCAACATATGCAAGATTTGAAGCCGAG 1560
Db 1501 GAGGTATGTTCAAGGTTCTCTTTATGCTGCTGCAACATATGCAAGATTTGAAGCCGAG 1560
QY 1561 CATGTGAAGCTAGTGTGCTATGACGAGGACTGCAAGAAATTTTCTCCGCTATACAGT 1620
Db 1561 CATGTGAAGCTAGTGTGCTATGACGAGGACTGCAAGAAATTTTCTCCGCTATACAGT 1620
QY 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATGATGATCTGTTG 1680
Db 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATGATGATCTGTTG 1680
QY 1681 AACATGTATGCGCTGATATGATGAGCTGCTGTTTATGCAAGAGCTGTAAGACCTCT 1740
Db 1681 AACATGTATGCGCTGATATGATGAGCTGCTGTTTATGCAAGAGCTGTAAGACCTCT 1740
QY 1741 GAAAACTTTGAAACTATATGATTTTACATTTGACAGCTGGGCTCTCAGAGAGTACGTTGAT 1800
Db 1741 GAAAACTTTGAAACTATATGATTTTACATTTGACAGCTGGGCTCTCAGAGAGTACGTTGAT 1800
QY 1801 GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGTAAAGTCTTA 1860
Db 1801 GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGTAAAGTCTTA 1860
QY 1861 GCTGCTGCTGCTGCAATTTGAGCTGATTTCACTGTTTACGCCAAGATTTTCTTAAAGC 1920
Db 1861 GCTGCTGCTGCTGCAATTTGAGCTGATTTCACTGTTTACGCCAAGATTTTCTTAAAGC 1920
QY 1921 AGCTACTTTTCAAGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
Db 1921 AGCTACTTTTCAAGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
QY 1981 GGGTCACTGAGAGCTGACATTCAGAAAGCACTTCCAGATGAGTCTAGACTGCAAG 2040
Db 1981 GGGTCACTGAGAGCTGACATTCAGAAAGCACTTCCAGATGAGTCTAGACTGCAAG 2040
QY 2041 AATATATGATTCAGAGTGGCAAGAAATTAAGTCTTGGCTTTTGGGCTGATCACCGCAT 2100
Db 2041 AATATATGATTCAGAGTGGCAAGAAATTAAGTCTTGGCTTTTGGGCTGATCACCGCAT 2100
QY 2101 GAAATTTTACAGAGGTTTGAATGGGCGAATGCTGAAGTTTGACTGACAGAGAGCT 2160
Db 2101 GAAATTTTACAGAGGTTTGAATGGGCGAATGCTGAAGTTTGACTGACAGAGAGCT 2160
QY 2161 GAAACTGCGCAGCTGGGTTGTTTATGATTTATACATGTTGAACATATCTGTTGA 2220

Db 474 TTCAGGACGACGCTTTAATCAGCCGAGACAGATTCTTCAGACTGCTTCGAAACCTTG 533
 Oy 421 TCTAACTCCGCTCTGAAGAGAGTCAATGAAGCTCTTGATATGAAGAAGCTACA 480
 Db 534 TCTAACTCCGCTCTGAAGAGAGTCAATGAAGCTCTTGATATGAAGAAGCTACA 593
 Oy 481 GTCATCATGATATGCTCTGAGTAAGGTTCTGCGGCTCTCTGTATATGCAAGAAGT 540
 Db 594 GTCATCATGATATGCTCTGAGTAAGGTTCTGCGGCTCTCTGTATATGCAAGAAGT 653
 Oy 541 GGTGAGACTGAGATAGTTCTTCGCGGTTGTGAGGCTCTGCTTAAAGAGAGCTTG 600
 Db 654 GGTGAGACTGAGATAGTTCTTCGCGGTTGTGAGGCTCTGCTTAAAGAGAGCTTG 713
 Oy 601 TCGTTTAAAGCAATATGCTTTAGTATAGCCGCTTGCTTCGATATGCTGAGGAGT 660
 Db 714 TCGTTTAAAGCAATATGCTTTAGTATAGCCGCTTGCTTCGATATGCTGAGGAGT 773
 Oy 661 GCTATGACATTTGATCCACTGATTTTATAGTATAGTATAGTATGAGAAAGCTTG 720
 Db 774 GCTATGACATTTGATCCACTGATTTTATAGTATAGTATAGTATGAGAAAGCTTG 833
 Oy 721 AAGCTTTTACAGAGAGAGAGCAAGTATGCTTGACCCGATTTACGTGCAAAATTGAT 780
 Db 834 AAGCTTTTACAGAGAGAGAGCAAGTATGCTTGACCCGATTTACGTGCAAAATTGAT 893
 Oy 781 GAGACTTTGAGAAAGATCACTCCGCTTATGCTTGAGAGTACTTGCTTAAAGCTTG 840
 Db 894 GAGACTTTGAGAAAGATCACTCCGCTTATGCTTGAGAGTACTTGCTTAAAGCTTG 953
 Oy 841 GATGATTAACGCTCGAAGAAAGTAAATGTTTAAAGCGGTGCGAATATTTTGTGCT 900
 Db 954 GATGATTAACGCTCGAAGAAAGTAAATGTTTAAAGCGGTGCGAATATTTTGTGCT 1013
 Oy 901 GTTGAAGAGGTGAGAGATCAGCTCTTGTGCGGTTTGAACCGGTGAAGATTATGAT 960
 Db 1014 GTTGAAGAGGTGAGAGATCAGCTCTTGTGCGGTTTGAACCGGTGAAGATTATGAT 1073
 Oy 961 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCAGC 1020
 Db 1074 GAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCAGC 1133
 Oy 1021 AATATTCAGCAGAGTATTTGAAGTTTGAAGTTGACAGTGTGCTTGTGCTCAAGT 1080
 Db 1134 AATATTCAGCAGAGTATTTGAAGTTTGAAGTTGACAGTGTGCTTGTGCTCAAGT 1193
 Oy 1081 TTTATTTGTAAGAGCACAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTT 1140
 Db 1194 TTTATTTGTAAGAGCACAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTT 1253
 Oy 1141 CAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGTATGATTAACGGAATAT 1200
 Db 1254 CAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGTATGATTAACGGAATAT 1313
 Oy 1201 GAGATTAAGCTTCGCTTAAGAAAGGAGCTGCTGCACTGCTTATAGGCAAGTTGAT 1260
 Db 1314 GAGATTAAGCTTCGCTTAAGAAAGGAGCTGCTGCACTGCTTATAGGCAAGTTGAT 1373
 Oy 1261 TGCCGTATGCTTGGCTTGAACAGTGAAGATTCACAAATTAAGAAATTCAGCTAT 1320
 Db 1374 TGCCGTATGCTTGGCTTGAACAGTGAAGATTCACAAATTAAGAAATTCAGCTAT 1433
 Oy 1321 GAGTTTGTGTAAGAAATTCAAATGCTGATGACAAATGATCTCCCTGACATATG 1380
 Db 1434 GAGTTTGTGTAAGAAATTCAAATGCTGATGACAAATGATCTCCCTGACATATG 1493
 Oy 1381 TTGTTGAAACCTGCTTGGCAGGAGTGTCTTTCTAGCTTCAAGACACCAAAAGAT 1440
 Db 1494 TTGTTGAAACCTGCTTGGCAGGAGTGTCTTTCTAGCTTCAAGACACCAAAAGAT 1553
 Oy 1441 AAATTTAACTCGGAGCTATGATGATCTTATGCTTGTGAGTTCTTGAAGAGT 1500

Db 1554 AAATTTAACTCGGAGCTATGATGATTCCTATGCTTTTGAAGTTACTTGAAGAGT 1613
 Oy 1501 GAGTATGTTCAAGGTTCTCTCTTATGCTGCTGCAACTATGAGCAAGATTGAGCCGAG 1560
 Db 1614 GAGTATGTTCAAGGTTCTCTCTTATGCTGCTGCAACTATGAGCAAGATTGAGCCGAG 1673
 Oy 1561 CATGTGAAGCTATGCTTATGACAGGACCTGCAAGAAAGTTTCTTCCGCTATACAGT 1620
 Db 1674 CATGTGAAGCTATGCTTATGACAGGACCTGCAAGAAAGTTTCTTCCGCTATACAGT 1733
 Oy 1621 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATAGTATGATCCTGTTGAT 1680
 Db 1734 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATAGTATGATCCTGTTGAT 1793
 Oy 1681 AACATGATGAGCCGATGATGATGAGCTGCTGCTTATATGCAAGCTGTAAGACCTCT 1740
 Db 1794 AACATGATGAGCCGATGATGATGAGCTGCTGCTTATATGCAAGCTGTAAGACCTCT 1853
 Oy 1741 GAAACTTTGAAACTATGATATGCAATTCGAGCTGGGCTCTCAGAGATGAGCTTGAT 1800
 Db 1854 GAAACTTTGAAACTATGATATGCAATTCGAGCTGGGCTCTCAGAGATGAGCTTGAT 1913
 Oy 1801 GAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA 1860
 Db 1914 GAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA 1973
 Oy 1861 GCTGCTGATGAGCAATTTGAGTATGATTTCACTGTTGAGCAGAGATTTTCTTAAAG 1920
 Db 1974 GCTGCTGATGAGCAATTTGAGTATGATTTCACTGTTGAGCAGAGATTTTCTTAAAG 2033
 Oy 1921 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATAGAAATCTGATGCTGATACAT 1980
 Db 2034 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATAGAAATCTGATGCTGATACAT 2093
 Oy 1981 GGTGCTGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 2094 GGTGCTGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2153
 Oy 2041 AATATGATTCGAAGTGAAGAAAGTATGATGATGATGATGATGATGATGATGAT 2100
 Db 2154 AATATGATTCGAAGTGAAGAAAGTATGATGATGATGATGATGATGATGATGAT 2213
 Oy 2101 GAAATGTTACAGAGGTTTGTGATGAGGCAATGCTGAAGATTTGACTGACAGAGCT 2160
 Db 2214 GAAATGTTACAGAGGTTTGTGATGAGGCAATGCTGAAGATTTGACTGACAGAGCT 2273
 Oy 2161 GAACTGCGCAGCTTGGGTTGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2274 GAACTGCGCAGCTTGGGTTGATGATGATGATGATGATGATGATGATGATGAT 2333
 Oy 2221 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGAGCAACTCTGAGAGAGTCTGCT 2280
 Db 2334 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGAGCAACTCTGAGAGAGTCTGCT 2293
 Oy 2281 TGTCTATCTGATTTGCTTATCAGAAACCAATGCTACTGATGCTGCAACCTACAC 2340
 Db 2394 TGTCTATCTGATTTGCTTATCAGAAACCAATGCTACTGATGCTGCAACCTACAC 2453
 Oy 2341 AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 Db 2454 AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2513
 Oy 2401 TCATTA 2406
 Db 2514 TCATTA 2519

RESULT 4
 AY221468 3668 bp DNA linear pln 05-AUG-2003
 LOCUS
 DEFINITION
 Arabidopsis thaliana division protein (AR6) gene, complete cds;
 nuclear gene for chloroplast product.
 ACCESSION
 AY221468

```

VERSION      AY221468.1  GI:33436338
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS      1 (bases 1 to 3668)
               Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Erp,H. and
               Osteryoung,K.W.
TITLE         ARC6 Is a U-Domain Plastid Division Protein and an Evolutionary
JOURNAL       Descendant of the Cyanobacterial Cell Division Protein Ftn2
PUBMED        Plant Cell 15 (8), 1918-1933 (2003)
12897262
AUTHORS      2 (bases 1 to 3668)
               Vitha,S., Koksharova,O., van Erp,H., Froehlich,J.E. and
               Osteryoung,K.W.
DIRECT SUBMISSION
JOURNAL       Submitted (17-JAN-2003) Department of Plant Biology, Michigan State
UNIVERSITY    University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
SOURCE
               1..3668
               /organism="Arabidopsis thaliana"
               /mol_type="genomic DNA"
               /db_xref="taxon:3702"
               /ecotype="Wassilewskija"
               478..3515
               /gene="ARC6"
               /note="synonym: At5g42480"
               join(<478..984,1067..1291,1379..1639,1723..2710,
               3000..3133,3225..3515)
               /gene="ARC6"
               /product="division protein"
               join(478..984,1067..1291,1379..1639,1723..2710,3000..3133,
               3225..3515)
               /gene="ARC6"
               /note="inner envelope membrane-localized; similar to
               cyanobacterial cell division protein slt0169 of
               Synechocystis; contains U-domain"
               /codon_start=1
               /product="division protein"
               /protein_id="AA018645.1"
               /db_xref="GI:33436339"
               /translation="MEALSHVIGTSLSPOLCRUPATTKLRSHSTSTTICASKYAD
               RILSDPNTSSSSSFATRTTATVSLDPSIDRBRHPRIPIIDYQVGAQTHLT
               DGRRAFEARVSKPQPFESDIALISRKQLQAACETLSPKSRKRNIBELDEBAT
               VITDPMKRVGALCVLQEGETEIVLRGELLKRLPKSFQDVLVWALPLAYS
               RDMALDPDPIFGVFEFEBALKLQEBGSLAPDLRAQIDTELEITPRVYLELG
               LPLGDDYAAKRLNGSGVRNIWVGSGGASLVGLTRKPMNBAFLMTAAEOVL
               FVATPSNIPAESFEVYEVALVAQAQFICKPHLLDADKPOQLQAQVMMEIRAM
               LNDTRNMEIDFLEGLCALLIGKVDCEKMLGLDSEDSQIRNPALIVEFVLENSRD
               DNDLPGLCLLETWLAGVFPFRFDTKDKKFLGDYDDPMVLSLEREVVQSGPL
               AAATMARIGAEHVASAMQALOKVPSRTDRNSABPKOVQTVSVFVGNVNRD
               GEGGVFLAEVRSENFTNDVAIRAGVSSSVDETVEVMADMKELKASVKILAGV
               AIGLILFISOKYIKSSSFQRKDWMSVDVATIGSVADSEALPRMDARTANI
               VSMWKIKSLARGPDRITMLPRVLDGRMKITWTDRAETRAQGLVYDITLKLSDS
               VTYSADGTRALVEATLEESACLSDLVHPENNATDVRITTRIEVFMSKGMKITBSGV
               LMS"
ORIGIN
Query Match      71.7%; Score 1724; DB 15; Length 3668;
Best local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY      1  ATGAAGCTCTGAGTCAGTCGTCGTCCTCCCATTCGAATTATGCGATTACCA 60
      DB      478  ATGAAGCTCTGAGTCAGTCGTCGTCCTCCCATTCGAATTATGCGATTACCA 537
QY      61  CCGGCGACGCAAAAGCTCCGAGTGGCAGACCACTCTTCAACTATCTGCTCCGACG 120
      DB      538  CCGGCGACGCAAAAGCTCCGAGTGGCAGACCACTCTTCAACTATCTGCTCCGACG 597
QY      121  AATGGGCGACGCTTCTCTCGCACTTCAATTTCACCTCGATTCTCTCTCTCC 180

```

```

DB      598  AATGGGCGACGCTTCTCTCGCACTTCAATTTCACCTCGATTCTCTCTCTCC 657
QY      181  TTGGCCACCGCCACACACACCGGCACTCTGCTCTCTGACACATTTATGATCGTCC 240
      DB      658  TTGGCCACCGCCACACACACCGGCACTCTGCTCTCTGACACATTTATGATCGTCC 717
QY      241  GAAAGCGACGTCGCCATGCCATTTGATTTCTACAGGATTTAGAGGCTGAACCATTTCC 300
      DB      718  GAAAGCGACGTCGCCATGCCATTTGATTTCTACAGGATTTAGAGGCTGAACCATTTCC 777
QY      301  TTAACCGATGGAATTCAGAGAGCATTCGAGCTAAGGCTTTGAAACCGCGCAATTCGT 360
      DB      778  TTAACCGATGGAATTCAGAGAGCATTCGAGCTAAGGCTTTGAAACCGCGCAATTCGT 837
QY      361  TTCAAGGACGACGCTTTATCAGCCGAGACAGATTTCTTAAGCTGTTCCGAACTCTG 420
      DB      838  TTCAAGGACGACGCTTTATCAGCCGAGACAGATTTCTTAAGCTGTTCCGAACTCTG 897
QY      421  TCTAATCCTCGGCTAAGAGAGTCAATGAAGGCTTCTTGATGATGAAGAGCTTACA 480
      DB      898  TCTAATCCTCGGCTAAGAGAGTCAATGAAGGCTTCTTGATGATGAAGAGCTTACA 957
QY      481  GTTCATCACTGATGTTCTTGGGAT-----AGGTTCTTGCGGC 518
      DB      958  GTTCATCACTGATGTTCTTGGGATTAAGTATTCGATTTCCGAAATTAAGTTCTTTC 1017
QY      505  -----AGGTTCTTGCGGC 518
      DB      1018  GTTTTATTTATTCGAATTTGATTAAGAGAACTTTTATCTAATGAAGTTCTTGCGGC 1077
QY      519  TCTCTGTGATTTGCAAGAGTGTGTGACCTGATAGTTCTTGGGTTGTGAGGCTCT 578
      DB      1078  TCTCTGTGATTTGCAAGAGTGTGTGACCTGATAGTTCTTGGGTTGTGAGGCTCT 1137
QY      579  GCTTAAGAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTGTTAGTTATGCGCTTGC 638
      DB      1138  GCTTAAGAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTGTTAGTTATGCGCTTGC 1197
QY      639  GTTTCGTGATGTCGAGGAGTCTATGCGATTGATTCACCTGATTTTATCTGTGTA 698
      DB      1198  GTTTCGTGATGTCGAGGAGTCTATGCGATTGATTCACCTGATTTTATCTGTGTA 1257
QY      699  TGAAGTTGTGAGGAGCTTTGAAGCTTTAC----- 730
      DB      1258  TGAAGTTGTGAGGAGCTTTGAAGCTTTTACAGTAGTTGACTTGTGTAATTG 1317
QY      731  -----A 731
      DB      1318  ACGAGCGTTGGCTTTAAGAACTTTCTGATTTGATTACTTTGTTATGACTTGTGTA 1377
QY      732  GAGAGAGAGACAAAGTACCTTGACCGAGATTTACGTGACAAATGATGAGACTTTGGA 791
      DB      1378  GAGAGAGAGACAAAGTACCTTGACCGAGATTTACGTGACAAATGATGAGACTTTGGA 1437
QY      792  AAGATCATCTCCGCTTATGCTTGTGAGCTACTTGCTTACCGCTTGTGATGATACGC 851
      DB      1438  AAGATCATCTCCGCTTATGCTTGTGAGCTACTTGCTTACCGCTTGTGATGATACGC 1497
QY      852  TGGGAAAGACCTAAATGTTTAAAGCGTGTGCGGAATTTTGTGCTGTTGAGAGAG 911
      DB      1498  TGGGAAAGACCTAAATGTTTAAAGCGTGTGCGGAATTTTGTGCTGTTGAGAGAG 1557
QY      912  TGGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTGATGAGCGCTTTT 971
      DB      1558  TGGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTGATGAGCGCTTTT 1617
QY      972  ACGAATGACAGCTGCTGAGC----- 991
      DB      1618  ACGAATGACAGCTGCTGAGCAGGTATACAGTTAGATACCTTTTAAATTTCTTTAGCA 1677
QY      992  -----AGGTTATCTTTTGA 1008

```

Db	1678	TGATATTAACCTTTAGCTTTCATATTTTAATGATATGTGTGTAGTTGAATCTTTTGTGA	1737
Qy	1009	GCTACCCCAAGCAATATTTCCAGCAGAGTCATTTTGAAGTTTACGAAGTTGCACCTTCTCTT	1068
Db	1738	GCTACCCCAAGCAATATTTCCAGCAGAGTCATTTTGAAGTTTACGAAGTTGCACCTTCTCTT	1797
Qy	1069	GTGGCTCAAGCTTTTATTTGTATTAAGGCAACCTTTTACAGATNGCTGATTAAGCAATTC	1128
Db	1798	GTGGCTCAAGCTTTTATTTGTATTAAGGCAACCTTTTACAGATNGCTGATTAAGCAATTC	1857
Qy	1129	CAGCAACTTCAGCAGGCTAAGATATGCTATGAGATTTCCCTGCAGATTTGTATGATATCA	1188
Db	1858	CAGCAACTTCAGCAGGCTAAGATATGCTATGAGATTTCCCTGCAGATTTGTATGATATCA	1917
Qy	1189	CGAATAATTTGGAGATATGACTTGGCTGTAGAAAAGGAGCTCTGTGCACCTGCTATAGGC	1248
Db	1918	CGAATAATTTGGAGATATGACTTGGCTGTAGAAAAGGAGACTCTGTGCACCTGCTATAGGC	1977
Qy	1249	AAAGTTGATGAAATGCCATATGTGTGTGGGCTTTAACAATGAGATTCACATATATAGAAAT	1308
Db	1978	AAAGTTGATGAAATGCCATATGTGTGTGGGCTTTAACAATGAGATTCACATATATAGAAAT	2037
Qy	1309	CCAGCTATTTGGAGTTGTTTGTGGAGAAATTCAAATCTGTATGCAATATGATATCTCCCT	1368
Db	2038	CCAGCTATTTGGAGTTGTTTGTGGAGAAATTCAAATCTGTATGCAATATGATATCTCCCT	2097
Qy	1369	GGACTATGCAAAATTTGTTGAAAACCTGTGTGGCAAGGGGTGTCTTCTTCTAGGTTCAAGAC	1428
Db	2098	GGACTATGCAAAATTTGTTGAAAACCTGTGTGGCAAGGGGTGTCTTCTTCTAGGTTCAAGAC	2157
Qy	1429	ACCAAGATTAATAAATTTTAAACTCGGGGACATATATATATCCTATAGGTTTGAATTAC	1488
Db	2158	ACCAAGATTAATAAATTTTAAACTCGGGGACATATATATATCCTATAGGTTTGAATTAC	2217
Qy	1489	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTCTCCTTATAGCTGCTGCACATATGGCAAG	1548
Db	2218	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTCTCCTTATAGCTGCTGCACATATGGCAAG	2277
Qy	1549	ATTGAGCCGAGCATGTGAAAAGCTATGATCTATGACAGCATGTCAGAAAAGTTTTCTTCC	1608
Db	2278	ATTGAGCCGAGCATGTGAAAAGCTATGATCTATGACAGCATGTCAGAAAAGTTTTCTTCC	2337
Qy	1609	CGCTATACAGATAGAAATCTCGGCTGTAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA	1668
Db	2338	CGCTATACAGATAGAAATCTCGGCTGTAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA	2397
Qy	1669	GATCCGTGTGGTAACAATGTAGGCCCGTATGCTGATGCTGATATGTTAAAGAGGCAAGCT	1728
Db	2398	GATCCGTGTGGTAACAATGTAGGCCCGTATGCTGATGCTGATATGTTAAAGAGGCAAGCT	2457
Qy	1729	GTAAAGACCTCTGAAAACCTTTGAAACTATATGATTTATGCAATTCGAGCTGGGGTCTCAGAG	1788
Db	2458	GTAAAGACCTCTGAAAACCTTTGAAACTATATGATTTATGCAATTCGAGCTGGGGTCTCAGAG	2517
Qy	1789	AGTAGCGTTGATGAAACTATCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	1848
Db	2518	AGTAGCGTTGATGAAACTATCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	2577
Qy	1849	GTGAAGATCCCTAGCGCTGCTGTGTGGCAATTTGAGCATTTTCACTGTGACGAGAAAGAT	1908
Db	2578	GTGAAGATCCCTAGCGCTGCTGTGTGGCAATTTGAGCATTTTCACTGTGACGAGAAAGAT	2637
Qy	1909	TTTCTTAATAAGCAGCTCATCTTTTTCAGAGCAAGATATGTTTCTTCTATGGAATCTGAT	1968
Db	2638	TTTCTTAATAAGCAGCTCATCTTTTTCAGAGCAAGATATGTTTCTTCTATGGAATCTGAT	2697
Qy	1969	GTGCGTACCA-----	1978
Db	2698	GTGCGTACCAATGATGATTTAAATGATGCAATTTTCAATATCTGCAATGCTCAAAATA	2757
Qy	1979	-----	1978
Db	2758	TGCTTGTGTTGTAGACTAGAAACATAGTTCCCATTTAATACATGTCCCAAAAGTTGTAC	2817

OY	1979	-----	1978
Db	2818	AAGATTAAACAATGTTCTGTAGTAATAATTCACATAATATGCTGGAAATTTTTGATCAAA	2877
OY	1979	-----	1978
Db	2878	CTGTAGACAGAAAATGTAAATTTCACTCTCAACATTTTCTGTTTAGAATAACGTAGATTAAG	2937
OY	1979	-----	1979
Db	2938	AGATTGCCTTAGTGCTGTGCTTGTCACAATTTCTTCCCTGATTTTTCTTTTCGATTT	2997
OY	1980	AGCGTCACTCAGACTGAACCATTTGAGNAGCACTTCCAGANAATGGAATGCTAGACGTGAGA	2039
Db	2998	AGGGTCAGTCAGACCTGACGATTTAGNAGCACTTCCAGAAATGGAATGCTAGACGTGAGA	3057
OY	2040	GAATATAGTATCCAAAGTGGCAGAAAGATTAAATGCTCTGGCTTTTGGGCTGTATCACGCCAT	2099
Db	3058	GAAATATAGTATCCAAAGTGGCAGAAAGATTAAATGCTCTGGCTTTTGGGCTGTATCACGCCAT	3117
OY	2100	AGAAATGTTACCAG-----	2113
Db	3118	AGAAATGTTACCAGAGGTGAGGGAATTAATCTAACATTCATCAATTTGTGTGAACCTGT	3177
OY	2114	-----	-AGCTTTTGGATGGGC 2128
Db	3178	TGACATGATTAATATGCTGTGCTGCTTGTGTTGATTTCTGTATTAATAGATTTTGGATGGGC	3237
OY	2129	GAATGCTGAAGATTTGGACTGACAGACAGCTGAAACCTGGCAGCTTGGGCTGTTATG	2188
Db	3238	GAAATGCTGAAGATTTGGACTGACAGACAGCTGAAACCTGGCAGCTTGGGCTGTTATG	3297
OY	2189	ATTATACACTGTGAAAATCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCGGTG	2248
Db	3298	ATTATACACTGTGAAAATCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCGGTG	3357
OY	2249	CTCTGGTGGAGCAACTCTGGAGGAGTCTGCTGTCTATCTGATTTGGTTTCATCCAGAAA	2308
Db	3358	CTCTGGTGGAGCAACTCTGGAGGAGTCTGCTGTCTATCTGATTTGGTTTCATCCAGAAA	3417
OY	2309	ACAATGCTACTGATGTGCAAGACTTACACACACACAGATACGAAGTTTTCGTCCAGCTCAG	2368
Db	3418	ACAATGCTACTGATGTGCAAGACTTACACACACACAGATACGAAGTTTTCGTCCAGCTCAG	3477
OY	2359	GGTGAAATCACTGGAAGGCTCTGTCTTCGTGCATCATTA	2406
Db	3478	GGTGAAATCACTGGAAGGCTCTGTCTTCGTGCATCATTA	3515
<hr/>			
RESULT 5			
LOCUS	AY221467	3664 bp	DNA linear
DEFINITION	Arabidopsis thaliana truncated division protein (arc6) gene, arc6-1 allele, complete cds; nuclear gene for chloroplast product.		
ACCESSION	AY221467		
VERSION	AY221467.1	GI:33436274	
KEYWORDS	.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsie.		
REFERENCE	Vitsha,S., Froehlich,J.E., Kokeharova,O., Pyke,K.A., Van Erp,H. and Oosteryoung,K.W.		
AUTHORS	Arc6 is a U-domain Placidal Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2 Plant Cell 15 (8), 1918-1933 (2003)		
TITLE	JOURNAL PUBMED 12897262		
REFERENCE	2 (bases 1 to 3664)		
AUTHORS	Vitsha,S., Kokeharova,O., van Erp,H., Froehlich,J.E. and Oosteryoung,K.W.		

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES Location/Qualifiers

Source

1. 3664
/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/ecotype="Ws-2"

/note="genotype: arc6 mutant"

1. 3664

/gene="arc6"

/allele="arc6-1"

join(<478..984,1067..1291,1379..>.1621)

/gene="arc6"

/allele="arc6-1"

/product="truncated division protein"

join(478..984,1067..1291,1379..1621)

/gene="arc6"

/allele="arc6-1"

/note="truncation caused by premature stop codon"

/codon_start=1

/product="truncated division protein"

/protein_id="AAQ18644.1"

/db_xref="GI:33436275"

/translation="MEALSHVIGLSPFQCRLPATYKLRSHNTSTCSAKYADRLSDPNTSDSSSPATRTTATVLSLPSIDREHVPITIDYQVAGTHRTDGIRAEARVSKRPQPFSDALISRQILOAETLSNPSRREBELDEEATVITVPMKVPKALCVLQEGETEIVLRVEALKEPLPSFQDVLVVALFLVDSRDMAALDPDPIFTEFEVRLKILQEGESLAPDLRAQIDTLEETITRYVLELLGLPIDDVYAARLNLGSEVRNLIWSVGGASALVGLTRKKFNMAFL"

ORIGIN

Query Match 71.5%; Score 1720.8; DB 15; Length 3664;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

QY 1 ATGAAGCTCTGATGTCACCTGCGCATTTGCTCTCCCATTCCTCAATATGCCATTAACA 60
DB 478 ATGAAGCTCTGATGTCACCTGCGCATTTGCTCTCCCATTCCTCAATATGCCATTAACA 537
QY 61 CCGGCGACGACAAAGCTCCGACGTAGGACCAACACCTCTACAACTATCTGCTCGGACG 120
DB 538 CCGGCGACGACAAAGCTCCGACGTAGGACCAACACCTCTACAACTATCTGCTCGGACG 597
QY 121 AAATGGCGCAGCGCTCTCTCTCCGACTTCAATTCACTCCGATTCCTCTCTCTCC 180
DB 598 AAATGGCGCAGCGCTCTCTCTCCGACTTCAATTCACTCCGATTCCTCTCTCTCC 657
QY 181 TTGGCCACCGGACCAACACCGGACCTCTGCTCTGCGACCATTTATGATCGTCC 240
DB 658 TTGGCCACCGGACCAACACCGGACCTCTGCTCTGCGACCATTTATGATCGTCC 717
QY 241 GAAAGCGACGTCCCATTCCTCATTTGATTTACACAGGTATTAAGAGCTCAACACATTTCC 300
DB 718 GAAAGCGACGTCCCATTCCTCATTTGATTTACACAGGTATTAAGAGCTCAACACATTTCC 777
QY 301 TTAACCGATGAATCAAGAAAGCAATTCGAGCTAGGCTTTGAAACCGCGCAATTCGCT 360
DB 778 TTAACCGATGAATCAAGAAAGCAATTCGAGCTAGGCTTTGAAACCGCGCAATTCGCT 837
QY 361 TTGAGGACGACGCTTTAATCAAGCGGACGACATTTTCAAGCTGTTGCGAACTCTG 420
DB 838 TTGAGGACGACGCTTTAATCAAGCGGACGACATTTTCAAGCTGTTGCGAACTCTG 897
QY 421 TCTAATCTCGGTCTGAAGAGAGTCAATGAAGGTCTTCTGATGATGAAGAGCTACA 480
DB 898 TCTAATCTCGGTCTGAAGAGAGTCAATGAAGGTCTTCTGATGATGAAGAGCTACA 957
QY 481 GTCATCACTGATTTCTTTGGAT----- 504
DB 958 GTCATCACTGATTTCTTTGGATTAAGTAAATTCGATTTCCGAAATATAAGTTCTTC 1017

QY 505 -----AAGTCTCTGCGGC 518
DB 1018 GTTTTAATTTCAATGATGATTAAGAGAACTTTATCTAGTGAAGTTCTCTGCGGC 1077
QY 519 TCTCTGTATTTCAAGAAAGTGTGAGACTGAGATATGTTCTTCCGCTGTGAGAGCTCT 578
DB 1078 TCTCTGTATTTCAAGAAAGTGTGAGACTGAGATATGTTCTTCCGCTGTGAGAGCTCT 1137
QY 579 GCTTAAGAGAGTGTGCTTAAGTCTGTTAAGCAAGATGTGTTTACTTATGCGCTTGC 638
DB 1138 GCTTAAGAGAGTGTGCTTAAGTCTGTTAAGCAAGATGTGTTTACTTATGCGCTTGC 1197
QY 639 GTTTCGATGTCTCGAGGAGTGTATGAGATGATCCAGCTGATTTATTAATGCTGCTA 698
DB 1198 GTTTCGATGTCTCGAGGAGTGTATGAGATGATCCAGCTGATTTATTAATGCTGCTA 1257
QY 699 TGAAGTTGTTAGAAAGCTTTGAAGCTTTTAC----- 730
DB 1258 TGAAGTTGTTAGAAAGCTTTGAAGCTTTTACAGGTATGTAAGCTTGTGTAATTTG 1317
QY 731 -----A 731
DB 1318 ACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATTAATTTGATGCTGTGTA 1377
QY 732 GAGAGAAAGAGCAAGTACCTTGACCGGATTTACGTCACAAATGATGAGACTTGA 791
DB 1378 GAGAGAAAGAGCAAGTACCTTGACCGGATTTACGTCACAAATGATGAGACTTGA 1437
QY 792 AGAGATCACTCCGCTGATGCTCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAGC 851
DB 1438 AGAGATCACTCCGCTGATGCTCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAGC 1497
QY 852 TCGAAAAGACTTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTGTGAGAGAG 911
DB 1498 TCGAAAAGACTTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTGTGAGAGAG 1557
QY 912 TGAGAGATCACTCTTGTGGGGTTTGAACCGGTGAAGTTATGATGAAGAGGCTTTT 971
DB 1558 TGAGAGATCACTCTTGTGGGGTTTGAACCGGTGAAGTTATGATGAAGAGGCTTTT 1617
QY 972 ACGAATGACAGCTGTGAGC----- 991
DB 1618 ACGAATGACAGCTGTGAGC----- 991
QY 992 -----AGTTGATCTTTTGT 1008
DB 1678 TGATATACTTTAGTTTCTCATTTTAATGATGATGTGTGATGATGATGATGATGAT 1737
QY 1009 GCTACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTAGAGTTGACCTTGCTT 1068
DB 1738 GCTACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTAGAGTTGACCTTGCTT 1797
QY 1069 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCAACCTTTTACAGAGTCTGATTAAGCAATTC 1128
DB 1798 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCAACCTTTTACAGAGTCTGATTAAGCAATTC 1857
QY 1129 CAGCAACTTCAGAGAGCTTAAGTATGCTATGAGATTTCTGCAATGTTGATTAACA 1188
DB 1858 CAGCAACTTCAGAGAGCTTAAGTATGCTATGAGATTTCTGCAATGTTGATTAACA 1917
QY 1189 CCGAATTAATTTGGAGATTAAGCTTGGCTTAAGAAAGGAGCTGTGCACTGCTTAAGGC 1248
DB 1918 CCGAATTAATTTGGAGATTAAGCTTGGCTTAAGAAAGGAGCTGTGCACTGCTTAAGGC 1977
QY 1249 AAGTTGATCAATGCGGTATGTTGTTGAGCTTGAAGAGATTCACAAATATTAAGAAAT 1308
DB 1978 AAGTTGATCAATGCGGTATGTTGTTGAGCTTGAAGAGATTCACAAATATTAAGAAAT 2037
QY 1309 CCAAGTATTTGAGAGTTGTTTGAAGATTCAAATCGATGATGATGATGATGATGATGAT 1368
DB 2038 CCAAGTATTTGAGAGTTGTTTGAAGATTCAAATCGATGATGATGATGATGATGATGATGAT 2097
QY 1369 GCACTATGCAAAATTTGTTGAAGACTGTGTTGAGAGGGTGTCTTCTAGGTTCAAGAGAC 1428

exon
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MDH9"
/clone_1b="Miteui P1"
/ecotype="Columbia"
complement (59. .239)
/note="CDS is reported in Acc# AB023032
gene_id:KSJ14.12
unknown protein"
/number=1

CDS
/evidence=not experimental
complement (join(2056. .2079,2278. .2393,2991. .3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pit||T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. .6138,6462. .6595)
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1

CDS
/evidence=not experimental
/protein_id="BAB10474.1"
/db_xref="GI:9759469"
/translation="MWDPEAEEDRKKSSNLSLAKNSVLPFERGRFYEGAKRE
RLRRKGEEAVGCVKGTYPYNGVNEHPNKRGRPAKESIKTVSVAVSSSIR
YSLRMAKERRKPPPLPLNVAVSAMKSVTTERKRLSLCTACTSSGLAFLRLGS
EPGVHAKRTLPFSAV"
complement (join(6868. .7911,7989. .8414,8555. .8671,
8754. .9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="arm repeat containing protein"
/protein_id="BAB10475.1"
/db_xref="GI:9759470"
/translation="MTDDEATGDAERDEEYTAEEPIHEDVDANVEIHEEYKEDD
DCEGLVDIVSIEFLDQINGYRRTQKECFNLVRKILIPILDEIRGESPSCKF
LNLRLKFLAKLLETCSGSKIYALDGETMTTRPSIYEKRLVLAAPDELM1
SGADKIDSLCKQKKAKRTDQDIELAVMMVPSKTPRNADSAIIRLAKKLE
LOTIDDKTEITIAIOSLIDQKGLNITKOHITIELNFKKQLQLEATDITVONIK
LITKSTSLIPHEIRICPTTEIMADPIYIACOTYKESIOKWPDAHCKTCTROE
DHLSAPNPKLIMQWCKRNPKIPEKVSPOSQOKOBYSLVEALSSOLEEQ
RSVQKMRILARENPERVLIAAGALPLVLQSLYPSGSIQENAVTTLNLSIDEN
KKLISNEGALPNIEILENGRREARNSAALFSLMDENKVTIGLSNGIPLVLDL
OHTLKGKDALTALFNLNSANKRAIDAGIVQPLNLTNKNLGMIDELSLILL
LASHPEGROALIGOLSFIEITLVEFIRGTPNKECATSVLLELSNNSFTLALQFV
YEIVETTSCTRAQRKANALIQLSKSEQT"

CDS
10888. .12579
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10476.1"
/db_xref="GI:9759471"

CDS
/translation="MMISEKPLGEESIRDLLEYLTVSRRLVKSQKLKKIKHTEV
VEDBEIARAVNCLISVGCRAVDGDFEDSSNKRSSASBEKGLMTICGTEETL
DCPSYGRERFMKNNRKYLDAGDQDRKHVLPDILCMCLMRPLTSLNHLVCK
KMSMANTQRFLOMRREGSFQTPMLPLPALKGCSSGDHGYDVSDKMRLETDL
KGRFMTSVTIIHEEIIYVGRSMDRNSFKSHRGILVSPSIKAMRKIASRHRASLPY
VGAETVSEFTQTKONRQDRRPHLSRVGSESDVDEPHRLSVRRONRSADQNGK
SHRLIRKDLRLNRSKRPVLAIGGTGFLPDSGELYDSATNTWSMOGLPMKF
GVVSCGICNGIIFYAISENDKLSGYDIERGFMTIOTSPIPRVHETYPGLVCSNHL
FMLSVSCDGDQIGRNRKAWKIMELDLYLTWESVHPADPMDMNTATVSDONI
LMGIEFKITFGQVLSFTYCDVUTLEBSMHSVNSQSLNLSCTKTITALLHL"
complement (15087. .16778)
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1

CDS
/evidence=not experimental
/protein_id="BAB10477.1"
/db_xref="GI:9759472"
/translation="MMISEKPLGEESIRDLLEYLTVSRRLVKSQKLKKIKHTEV
EDBEVARAVNCLISVGCRAVDGDFEDSLNKRSSASBEKGLMTICGTEETL
DCPSYGRERFMKNNRKYLDAGDQDRKHVLPDILCMCLMRPLTSLNHLVCK
KMSMANTQRFLOMRREGSFQTPMLPLPALKGCSSGDHGYDVSDKMRLETDL
KGRFMTSVTIIHEEIIYVGRSMDRNSFKSHRGILVSPSIKAMRKIASRHRASLPY
VGAETVSEFTQTKONRQDRRPHLSRVGSESDVDEPHRLSVRRONRSADQNGK
SHRLIRKDLRLNRSKRPVLAIGGTGFLPDSGELYDSATNTWSMOGLPMKF
GVVSCGICNGIIFYAISENDKLSGYDIERGFMTIOTSPIPRVHETYPGLVCSNHL
FMLSVSCDGDQIGRNRKAWKIMELDLYLTWESVHPADPMDMNTATVSDONI
LMGIEFKITFGQVLSFTYCDVUTLEBSMHSVNSQSLNLSCTKTITALLHL"
complement (join(18111. .18193,18281. .18535,18773. .19001,
19088. .19195,19307. .19447,19542. .19952,20129. .20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
sp|P42251"

CDS
/evidence=not experimental
/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFWCLSLTLTVIAAEPTIRIARFSCANOGAPQIMDA
INKEFDQLFIMLDGNTYGDIRRLPLTVIGKERTEGPMRNSRPFVSSSEKRLRYAKK
ANNGYRIORAKAVIGTMDHDYGLNDAGKEPFRKYINQKMLDPLMEPDSRRQA
GVYASYTGPNNRKYIYLDTRHDPDLSDSISIGDTQMDMLENLSGPRSITII
GSSVOYISNLSATTPGFYMSWGRPKRKLPLKIDTKRGVAFISGDVHREGIT
RDCSVGPIYIDTSSGLVOSVKKVPRPLRSIVRLFPYTPPTMRYINDCKFKSCT
YGOQNGAISIDNNAPVTIKLEIRVNGDTVIGTVNSLELDPOGSNSLADTTFKK
SORYCTLEIELPNTTRYRLAVLIYFTIYAVLMAHILGILGAVALITACVYKCRVD"
complement (21540. .22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"
/translation="MTLAKQKSLRLYKVSQRSSEKRNLEDESRSSNGSSS
LNVNEKTFYDINDASDGKISGEBELQSVSLGALSREVEEVKTSVYDDQAPID
FEERLKLMEBESDSEBRKELEAFGMVMEBEFTTAAISLRTTSLRLESCTVDAK
KVMIRGFDNDVDSLFDEFLNMR"

CDS
26332. .25679,25821. .25928,26039. .26227,26314. .26613,
26730. .26831,27169. .27303,27487. .27711,27919. .28122,
28213. .28380,28646. .28795,28889. .29026,29204. .29431,
29519. .29671,29784. .29910,30115. .30241,30554. .30674,
30763. .30843,30932. .31033,31112. .31262,31338. .31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="p1b11ybs1n"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFTVKVSPIILAPFNSGDSRRARYLKSNNKVFNS
SPILTPHYVVEAPSLIPYNGILMAOPNSHKGILKNIVSGKATGISLSQGNFCLT
CRNQGAGIRALPDAVDTRAPFSLSSSUTSSIRKSQIYNATLGPBEPRAAGTAPD
GIVAEKODDLPLPEIDSALEAFGLCELPFSKREGLDCTGARSNAVYDPPHVFPI
BAHMEHVVDSIDDEBEOGIAHMIIEHVAFSGSKREGLDCTGARSNAVYDPPHVFPI
HSPLTHGSEDDLPFVLDALNLSIAHPKLSRVKERRAIISELOMANTIEYRQD
OLLOHLSHKKLGRFPPIGLEBOIKKMDVKIRKFRHMYFANATLYIYGDINDPR
IYHNTIYAVKNGLNDNESPTSSPSGAFGAMNPLFLPKIAGIGTSPSKNTTQDS
KMKTERHAIRPPEVHNSIPGTSVULKFPQIKHLLONFALINMEKPIVSKQVTS
DLNVLNMRIFLSALHFRINTRKSNNPPTSVELHSDSGRGCTVTLTVIAEPON
MUNAIVVAVOEVRLEKFGVTRBELRYMDALIKDSBEHLAAMDVNSVMDNFIWES
DASHSTQTOHEITLVAAGVTLEEVTVGAKYLEFISDFGRTPALPAAIVCV
PTKPVHVGSESPENISPDIELSVNSGLAPLEAPBELEVEPGLSQGLKELTOR
NCPVAVIPSGELKADKXGTITOLNSGLIYVKKYTSPEAGVMBLGGGRABAR
TSDSKAAVVGATTSLEGRAVDPSRQBYELFVNNHINCSLESTEFIAMERPTLR
DNMQAFAVQLHNVLRSWLEDAFRAQULYSTRSIPKSLERKTAHKLAMGLNG
DERFVEPTPKSLQSLNLSYKDAVMHGFVDNNHVSIVGDFSEBIERCLIDVLGVYK
ASHDSKAPGSEBILFROPYAGIQFOQVFLKDDDEBAICAYIADPARNMFOFTVGDGL
FQSVSKLPVANDGLKSEBQLLEGDRELQKLRAPLFGVTMGLLAIINSLFTT

CDS

VRDLSGLTYDVSEFLNLPRLNLGMWTVISVTSIPGKYRAVADCKSVLRLSHNOIAP
RELDRAKRTILMKHEALSKSNAYMLNLHLQASVPRKLSICIKLVLSYKASLSD
IYLAVYQLHVEDSLYSICIGIAGAQBEEITVSEBEPEDSVSGVLPVGRSSMTTR
PTT"
complement (join(33251..33326,33512..33663,33744..33797,
33888..33968,34143..34265,34359..34768,34873..35052,

71.4%; Score 1717.6; DB 15; Length 85791;

Query Match Best Local Similarity 79.1%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

QY 1 ATGGAAGCTCTGATGACGTCGGCATTTGGTCTCTCCCATTTCAATTATGCCATTTACCA 60
Db 64077 ATGGAAGCTCTGATGACGTCGGCATTTGGTCTCTCCCATTTCAATTATGCCATTTACCA 64136
QY 61 CCGGCGACGCAAGAAAGCTCCGACGTAGACACACACCTCTACACTATCTGCTCCGACG 120
Db 64137 CCGGCGACGCAAGAAAGCTCCGACGTAGACACACACCTCTACACTATCTGCTCCGACG 64196
QY 121 AATGAGCGCAGCGCTCTCTCTCGACTTTCATTTTCACCTCGATTCTCTCTCTCC 180
Db 64197 AATGAGCGCAGCGCTCTCTCTCGACTTTCATTTTCACCTCGATTCTCTCTCTCC 64256
QY 181 TTGCGCACCGGCAACCAACGCGCACTCTGCTCTCTGCGACCATTTGATCGTCCC 240
Db 64257 TTGCGCACCGGCAACCAACGCGCACTCTGCTCTCTGCGACCATTTGATCGTCCC 64316
QY 241 GAAAGCGACGTCGCCATCCCAATTGATTTCTACAGGATTTAGAGAGTGAACAACATTC 300
Db 64317 GAAAGCGACGTCGCCATCCCAATTGATTTCTACAGGATTTAGAGAGTGAACAACATTC 64376
QY 301 TTAACGAGTGAATCAGAAAGACATTCGAGTAAAGGTTTCAAAACCGCGCAATTCGAT 360
Db 64377 TTAACGAGTGAATCAGAAAGACATTCGAGTAAAGGTTTCAAAACCGCGCAATTCGAT 64436
QY 361 TTCAGCGACGACGCTTTAATCAGCCGCGACACATTTCTCAAGCTGCTTCGAAACTCTG 420
Db 64437 TTCAGCGACGACGCTTTAATCAGCCGCGACACATTTCTCAAGCTGCTTCGAAACTCTG 64496
QY 421 TCTAATCTCGGCTCTAGAAAGAGTACAAAGAAAGTCTTCTGATGATGAAGAAAGCTACA 480
Db 64497 TCTAATCTCGGCTCTAGAAAGAGTACAAAGAAAGTCTTCTGATGATGAAGAAAGCTACA 64556
QY 481 GTCATCACTGATGCTCTTGGGAT----- 504
Db 64557 GTCATCACTGATGCTCTTGGGATAGGTAATTCGAAATTAAGTTCTTC 64616
QY 505 -----AAGTTCCTGGGAC 518
Db 64617 GTTTTAATTCATGAATTGGATAAAGAAAGAACTTTATCTAGTGAAGGTTCTCGGTGC 64676
QY 519 TCTCTGTGTATTTGCAAGAGGTGTGAGTGAATGTTCTTCGGGTGTGAGAGCTCT 578
Db 64677 TCTCTGTGTATTTGCAAGAGGTGTGAGTGAATGTTCTTCGGGTGTGAGAGCTCT 64736
QY 579 GCTTAAGAGAGGTGCTAAGTCTGTTAAGCAAGATGTGTTTAACTTAAGGCGCTGC 638
Db 64737 GCTTAAGAGAGGTGCTAAGTCTGTTAAGCAAGATGTGTTTAACTTAAGGCGCTGC 64796
QY 639 GTTTCCTGATGTCTGAGGAGTGTATGCAATGATCACTGATTTTATTAATCACTGCTTA 698
Db 64797 GTTTCCTGATGTCTGAGGAGTGTATGCAATGATCACTGATTTTATTAATCACTGCTTA 64856
QY 699 TGAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 730
Db 64857 TGAGTTTGTGAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGGTGAATTTG 64916
QY 731 -----A 731
Db 64917 ACAGAGCTTGGCTTTAAGAACTTTCTGATTTGATACCTTTGATTTGAGCTTGTGTA 64976
QY 732 GGAGGAAGAGCAAGTACGCTTGCACGGAATTTACGTGCAAAATGATGAGACTTTGGA 791
Db 732 GGAGGAAGAGCAAGTACGCTTGCACGGAATTTACGTGCAAAATGATGAGACTTTGGA 791

Db 64977 GGAGGAAGAGCAAGTACGCTTGCACGGAATTTACGTGCAAAATGATGAGACTTTGGA 65036
QY 792 AGAGTACACTCCGCGTTATGTCTGGAGCTACTTGCTTAACGCTTGGTATGATTAACG 851
Db 65037 AGAGTACACTCCGCGTTATGTCTGGAGCTACTTGCTTAACGCTTGGTATGATTAACG 65096
QY 852 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTCTGTGGAGAG 911
Db 65097 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTCTGTGGAGAG 65156
QY 912 TGGAGCATGACCTTGTGTGGGGGTTTGACCCGTGGAAGTTTATGAATGAGGCGTTT 971
Db 65157 TGGAGCATGACCTTGTGTGGGGGTTTGACCCGTGGAAGTTTATGAATGAGGCGTTT 65216
QY 972 ACGAATGACAGCTGTGAC----- 991
Db 65217 ACGAATGACAGCTGTGACAGGATATACGTTAGATACCTTTTAAATTCCTTAGCA 65276
QY 992 -----AGTTGATCTTTTGT 1008
Db 65277 TGATATACCTTATAGTTCATTTTATATGATGATGTGTGATGATGATCTTTTGT 65336
QY 1009 GCTAACCAGCAAGAAATTTCCAGAGAGTCAATTTGAAGTTTACAGAGTTGCACTTGCTTT 1068
Db 65337 GCTAACCAGCAAGAAATTTCCAGAGAGTCAATTTGAAGTTTACAGAGTTGCACTTGCTTT 65396
QY 1069 GTGGCTCAAGCTTTTATGTTAAGAACCACTTTTACAGAGTCTGATGAAGCAATTC 1128
Db 65397 GTGGCTCAAGCTTTTATGTTAAGAACCACTTTTACAGAGTCTGATGAAGCAATTC 65456
QY 1129 CAGCAATTCAGAGCTTAAGGTAATGCTATGAGATTCCTGCAATGTTTATGATTA 1188
Db 65457 CAGCAATTCAGAGCTTAAGGTAATGCTATGAGATTCCTGCAATGTTTATGATTA 65516
QY 1189 CGAAATTAATTTGGAGATTAACCTCGCTTGAAGAGGGGACTCTGTCACCTGTTATAGG 1248
Db 65517 CGAAATTAATTTGGAGATTAACCTCGCTTGAAGAGGGGACTCTGTCACCTGTTATAGG 65576
QY 1249 AAGTGTATGAAATGCGTATGTTGTTGAGCTTGAAGAGTGAAGATTCACATTAAGAA 1308
Db 65577 AAGTGTATGAAATGCGTATGTTGTTGAGCTTGAAGAGTGAAGATTCACATTAAGAA 65636
QY 1309 CCAGTATTTGTGAGATTTGTTTGAAGAAATTCATATCTGATGAATGATATCTCCCT 1368
Db 65637 CCAGTATTTGTGAGATTTGTTTGAAGAAATTCATATCTGATGAATGATATCTCCCT 65696
QY 1369 GGAATATGCAAAATTTGGAAGAACTGTTGGAGGGGTTGCTTCTAGGTTCAAGAG 1428
Db 65697 GGAATATGCAAAATTTGGAAGAACTGTTGGAGGGGTTGCTTCTAGGTTCAAGAG 65756
QY 1429 ACCAAGATTAATAAATTTAACTCGGGGACTACTATGATGATCTATGTTGAGTTAC 1488
Db 65757 ACCAAGATTAATAAATTTAACTCGGGGACTACTATGATGATCTATGTTGAGTTAC 65816
QY 1489 TTGAAAAGAGTGAAGTGAAGTTCAGGCTTCTCTTAACTGCTGCTGCACTATGCAAG 1548
Db 65817 TTGAAAAGAGTGAAGTGAAGTTCAGGCTTCTCTTAACTGCTGCTGCACTATGCAAG 65876
QY 1549 ATTGAGCGCGAGATGTAAGAACTAGTGCATGCAAGGCACTGAGAAAGTTTCTCTCC 1608
Db 65877 ATTGAGCGCGAGATGTAAGAACTAGTGCATGCAAGGCACTGAGAAAGTTTCTCTCC 65936
QY 1609 CGCTATACAGATGAAGAACTCGGCTGAACCAAGAGATGTGCAAGAGACAGTGTATGTA 1668
Db 65937 CGCTATACAGATGAAGAACTCGGCTGAACCAAGAGATGTGCAAGAGACAGTGTATGTA 65996
QY 1669 GATCTGTTGTGTAACAATGATGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGACT 1728
Db 65997 GATCTGTTGTGTAACAATGATGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGACT 66056
QY 1729 GTAAGACCTCTGAAAACCTTTGAACATAATGATTAAGCAATTCGAGCTGGGGCTCAGAG 1788
Db 66057 GTAAGACCTCTGAAAACCTTTGAACATAATGATTAAGCAATTCGAGCTGGGGCTCAGAG 66116

Dh 132882 ATTGACGTTCTGGTGCAATATGCTTTCATGATGACCTTGATCATGTAGAGAGAGG 132941
Qy 743 CAAGTACCTTGACCGGATTTACGTGCAAAATGTAGACTTTGGAAGATCATCTC 802
Db 132942 CAGACGCTTAGACCGGATTTACAAACAAATGTATGATGACTTGAAGATTAACCC 133001
Qy 803 CCGCTATGCTTGGAGCTACTTGGCTTACCGCTGGTGTATGATTAACGTCGCAAAAGAC 862
Db 133002 CAGCTGTGTATTAAGAACTTTAGCTTGCTCTTGTATGATGAACATCGAGCGCGAGAG 133061
Qy 863 TAAATGCTTAAAGCGGTGTCGGAAATTTGTGTCTGTGTGGAGAGAGTGAGCATCG 922
Db 133062 AAGAAAGTCTCCAAAGGTGTCGCAACATTCGTGGGAGTGAGGTGGAGGAGCAG 133121
Qy 923 CTCTGTGGGGTTTGACCCGAGAAATTTATGAATAGAGGCTTTTACGAATGACAG 982
Db 133122 CAATGCTGGAGATTTACACGTGAGATTTCAATGATAGGATTCCTGCATATGAAG 133181
Qy 983 CTGCTGAGC----- 991
Db 133182 CTGCCAAGAGGTTCATAGATATATGTGCAATTTTTTAATTTCTCCCG 133241
Qy 992 ----- 991
Db 133242 CCCCTTAGTTTTTTTATGAAGTTTACATTATCTCTTCATTAATCTTTTATC 133301
Qy 992 -----AGTGATCTTTTGTAGTACCCCAAGCAATTCACAGAG 1034
Db 133302 TCCAAATCTTTTTTTTAAAGGTGCAACTTTTGTAGCAACCTTAGCAATATTCAGCTGA 133361
Qy 1035 GTCATTTGAATTTAGCAAGTTGACCTGCTCTTGTGGCTCAAGCTTTTATTTGTAGAA 1094
Db 133362 AAGTTTGAAGCCTTAGGGGTGCACTTGATGTTGACAAAGCCTTTGTAGTAAGAA 133421
Qy 1095 GCCACACCTTTTACAGATGCTATTAAGCAATTCAGCAACTTCAGAGCCTTAGATAT 1154
Db 133422 GCCACATCTTATCAAGATGCTATTAATTTATTCATCAACTTCAACAACTAAGGTAAAC 133481
Qy 1155 GGGTATGAGATTCCTGCCATGTTGTATGATGAC-----ACGAATTAATTTGGAGATAGA 1208
Db 133482 AATATAGAGAAAGTCCCTCTGTATTACTCCCATGAGATGAGAGAGAGAGAGTTGA 133541
Qy 1209 CTTGCGCTAGAAAGGGAATCTGTGCACTGTTATAGCAAAAGTTGATGAATGCCGTAT 1268
Db 133542 TTTTGCATTAAGAAAGGGGTGTGTGCACTGCTTGTGGGAGCTTGATCAATGTGATC 133601
Qy 1269 GTGGTTGGCTTAGACAGTGAAGATTCAAAATATAGAAATCCACTATTGTGAGTTGT 1328
Db 133602 ATGGTTGGGTCTGATATGATACAGCTCACTTATCGAAACCCACTTATATAGACTTAT 133661
Qy 1329 TTTTGAAGATTAATTCGTATGATGACATGATGATCTCCCTGGACTATGCAATTTGTGA 1388
Db 133662 TATGGAAGAAAGCAAAAGGTATAGACAGTGAATCTTCGGACTGTAAATTTGTGGA 133721
Qy 1389 AACCTGGTTGGAGCGGTTGTCTTTCTAGATTCAAGACACCAAAAGATTAATAATTT 1446
Db 133722 GACATGTTGATGAGAGTGTCTTTCTAGGTTTAAGATATCTAAGAGCAAACTTT 133779

RESULT 8
AP008208 011/c
WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208
Fragment Name Begin End
AP008208_000 1 110000
AP008208_001 100001 210000
AP008208_002 200001 310000
AP008208_003 300001 410000
AP008208_004 400001 510000
AP008208_005 500001 610000
AP008208_006 600001 710000
AP008208_007 700001 810000
AP008208_008 800001 910000

AP008208_009 900001 1010000
AP008208_010 1000001 110000
AP008208_011 1100001 1210000
AP008208_012 1200001 1310000
AP008208_013 1300001 1410000
AP008208_014 1400001 1510000
AP008208_015 1500001 1610000
AP008208_016 1600001 1710000
AP008208_017 1700001 1810000
AP008208_018 1800001 1910000
AP008208_019 1900001 2010000
AP008208_020 2000001 2110000
AP008208_021 2100001 2210000
AP008208_022 2200001 2310000
AP008208_023 2300001 2410000
AP008208_024 2400001 2510000
AP008208_025 2500001 2610000
AP008208_026 2600001 2710000
AP008208_027 2700001 2810000
AP008208_028 2800001 2910000
AP008208_029 2900001 3010000
AP008208_030 3000001 3110000
AP008208_031 3100001 3210000
AP008208_032 3200001 3310000
AP008208_033 3300001 3410000
AP008208_034 3400001 3510000
AP008208_035 3500001 3610000
AP008208_036 3600001 3710000
AP008208_037 3700001 3810000
AP008208_038 3800001 3910000
AP008208_039 3900001 4010000
AP008208_040 4000001 4110000
AP008208_041 4100001 4210000
AP008208_042 4200001 4310000
AP008208_043 4300001 4410000
AP008208_044 4400001 4510000
AP008208_045 4500001 4610000
AP008208_046 4600001 4710000
AP008208_047 4700001 4810000
AP008208_048 4800001 4910000
AP008208_049 4900001 5010000
AP008208_050 5000001 5110000
AP008208_051 5100001 5210000
AP008208_052 5200001 5310000
AP008208_053 5300001 5410000
AP008208_054 5400001 5510000
AP008208_055 5500001 5610000
AP008208_056 5600001 5710000
AP008208_057 5700001 5810000
AP008208_058 5800001 5910000
AP008208_059 5900001 6010000
AP008208_060 6000001 6110000
AP008208_061 6100001 6210000
AP008208_062 6200001 6310000
AP008208_063 6300001 6410000
AP008208_064 6400001 6510000
AP008208_065 6500001 6610000
AP008208_066 6600001 6710000
AP008208_067 6700001 6810000
AP008208_068 6800001 6910000
AP008208_069 6900001 7010000
AP008208_070 7000001 7110000
AP008208_071 7100001 7210000
AP008208_072 7200001 7310000
AP008208_073 7300001 7410000
AP008208_074 7400001 7510000
AP008208_075 7500001 7610000
AP008208_076 7600001 7710000
AP008208_077 7700001 7810000
AP008208_078 7800001 7910000
AP008208_079 7900001 8010000
AP008208_080 8000001 8110000
AP008208_081 8100001 8210000

AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000
AP008208_132	13200001	13310000
AP008208_133	13300001	13410000
AP008208_134	13400001	13510000
AP008208_135	13500001	13610000
AP008208_136	13600001	13710000
AP008208_137	13700001	13810000
AP008208_138	13800001	13910000
AP008208_139	13900001	14010000
AP008208_140	14000001	14110000
AP008208_141	14100001	14210000
AP008208_142	14200001	14310000
AP008208_143	14300001	14410000
AP008208_144	14400001	14510000
AP008208_145	14500001	14610000
AP008208_146	14600001	14710000
AP008208_147	14700001	14810000
AP008208_148	14800001	14910000
AP008208_149	14900001	15010000
AP008208_150	15000001	15110000
AP008208_151	15100001	15210000
AP008208_152	15200001	15310000
AP008208_153	15300001	15410000
AP008208_154	15400001	15510000

Query Match	Best Local Similarity	Score	DB	Length
AP008208_155	1550001	15610000		
AP008208_156	1560001	15710000		
AP008208_157	1570001	15810000		
AP008208_158	1580001	15910000		
AP008208_159	1590001	16010000		
AP008208_160	1600001	16110000		
AP008208_161	1610001	16210000		
AP008208_162	1620001	16310000		
AP008208_163	1630001	16410000		
AP008208_164	1640001	16510000		
AP008208_165	1650001	16610000		
AP008208_166	1660001	16710000		
AP008208_167	1670001	16810000		
AP008208_168	1680001	16910000		
AP008208_169	1690001	17010000		
AP008208_170	1700001	17110000		
AP008208_171	1710001	17210000		
AP008208_172	1720001	17310000		
AP008208_173	1730001	17410000		
AP008208_174	1740001	17510000		
AP008208_175	1750001	17610000		
AP008208_176	1760001	17710000		
AP008208_177	1770001	17810000		
AP008208_178	1780001	17910000		
AP008208_179	1790001	18010000		
AP008208_180	1800001	18110000		
AP008208_181	1810001	18210000		
AP008208_182	1820001	18310000		
AP008208_183	1830001	18410000		
AP008208_184	1840001	18510000		
AP008208_185	1850001	18610000		
AP008208_186	1860001	18710000		
AP008208_187	1870001	18810000		
AP008208_188	1880001	18910000		
AP008208_189	1890001	19010000		
AP008208_190	1900001	19110000		
AP008208_191	1910001	19210000		
AP008208_192	1920001	19310000		
AP008208_193	1930001	19410000		
AP008208_194	1940001	19510000		
AP008208_195	1950001	19610000		
AP008208_196	1960001	19710000		

```

Db      80900 ATGATTTCTTTCAAAAACACGGAATAGCATTCCTCGTAATGTTGAATTTACAT 80841
Oy      1054 GTTGACATTTGCTCTTGTGGCTCAAGCTTTATTATGGTAAGAACCAACCTTTACAGAT 1113
Db      80840 GTAGACATTTGCACATGTCCTCCAGCAATTTATAGTAAAGGACCAATTCATCATGATG 80781
Oy      1114 GGTGATTAAGCAATTCACAGCACTTGACAGGGCTAAGGTAATGCGTATGAGATTCCTGCG 1173
Db      80780 GCGGATGATCTTTTGAACACTCCAGAACTTCA-----C 80745
Oy      1174 ATGTTGATGATACACGGAATTAATGGAGATAGACTTGGCTGTAGAAAGGCACTGT 1233
Db      80744 ATAGGTTCTCATTAAGCTTATATGATGATGAGACTTGATGAAAGGCATTCCTGC 80685
Oy      1234 GCACGCTTATAGGCAAGTTGATGAATGCCATATGTTGGCTTAACAGCTGAGAT 1293
Db      80684 TCATTGCTAGTCGGAGATGTTGACAGTCAAGAAATGTGCTTGAATTAATATGATGCT 80625
Oy      1234 TCACAAATATAGGAATACAGCTATGTTGAGAGTTTGTGTAAGAAATCAATC---GTGAT 1350
Db      80624 TCACCAATACAGAGACCCCAAAATTTCTAGAGTTTATTTGACCACTTACATCACTGAA 80565
Oy      1351 GACCAATGATGATCTCCCTGCACTATGCAAAATGTTGAAAACCTGTGCGACGGGTTGTC 1410
Db      80564 GGAATGATCTTCTTCCAGGGCTGTCAGCTTTGAGAGCTTGCTTATCTTTGAGGTT 80505
Oy      1411 TTTTCTAGCTTACAGACACCAAGATTAATAATTTAACTCGGGGACTCATATGATGAT 1470
Db      80504 TTTTCTAGGAGACAGATATCTCGGGGCACTGACATTTGAGAAATTAATCAATGATGAT 80445
Oy      1471 CCTATGCTTTTGTGACTTCTGAAAAGATGAGAGTATTCAGGGTCTCTCTTAGCTGCT 1530
Db      80444 CCAAGAGTTTAACTAAGCTTAAAGAGATGAGAGGCTGCTCTTCTCATTTGGCTGCT 80385
Oy      1531 GCTGCACTATGCGCAAGATTTGAGC 1556
Db      80384 GCTGCTGCTATGCAAACTTGCTGC 80359

RESULT 9
AP004885/c 150462 bp DNA linear PLN 15-SEP-2004
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION PAC clone: P0575F10.
ACCESSION  AP004885
VERSION     AP004885.3 GI:41053009
KEYWORDS
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
  TITLE     Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
  JOURNAL   Cloning: P0575F10
  PUBLISHED Only in Database (2002)
  AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
  TITLE     Direct Submision
  JOURNAL   Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program, Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail: tsasak@nias.affrc.go.jp, URL: http://rsgp.dna.affrc.go.jp/,
            Tel: 81-298-38-7441, Fax: 81-298-38-7468)
            On Jan 21, 2004 this sequence version replaced gi:38142429.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mtc.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerX
            (http://www.tigr.org/tdb/glimmerx/glmr_form.html), RiceHMM
            (http://rsgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), slmd

```

FEATURES

source

(<http://globin.cse.psu.edu/html/docs/slmd.html>), gap2 (<http://www.tigr.org/software/glimmerx/>), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP or DBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs representing the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0575F10 clone has an overlap with OJ1020_C02 (DBJ: AP004078) clone at 5' end and with P0482P12 (DBJ: AP005311) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rsgp.dna.affrc.go.jp/Genomeseq.html>.

location/Qualifiers

```

1..150462
  organism="Oryza sativa (japonica cultivar-group)"
  mol_type="genomic DNA"
  cultivar="Nipponbare"
  db_xref="taxon:39947"
  chromosome="2"
  clone="P0575F10"
  complement(4987..8112)
  /gene="P0575F10.1"
  complement(join(4987..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..8112))
  /gene="P0575F10.1"
  note="supported by full-length cDNA(s): AK106553"
  complement(join(5003..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..7876))
  /gene="P0575F10.1"
  note="supported by full-length cDNA(s): AK073059"
  complement(join(5315..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..7876))
  /gene="P0575F10.1"
  note="AKU070160(BE1534)
  contains full-length cDNA(s): AK106553,AK073059"
  /product="putative eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa"
  protein_id="BAD07941.1"
  /db_xref="GI:41053010"
  translation="MEDMBSEDFOPAVSPAKAPPKSKNADVEDVDYKESWEEEEKKXPPEVETKAPKSGKGAAGKAGQSTSEVVDLADPALKLEKQRLVEADFKSTTRLEKKGDSKSLDTPIKXSBDPAEYALIANGLRKYKSFHYMGLIKVMRLSMASIKGADAKDISSTAIATNKKIKAEKAAAGKKKGAGKKQKHINKDDPIFGKGFDPDEYDFM"
  complement(10186..13754)
  /gene="P0575F10.2"
  complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
  /gene="P0575F10.2"
  note="start and end point are not identified"
  complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
  /gene="P0575F10.2"
  note="contains EST(s): AU058423(E51136), AU189658(E51136)"
  /codon_start=1
  /product="plastid division protein precursor"
  protein_id="BAD07942.1"

```

CDS

gene

mRNA

[illegible]

Query Match	7.5%	Score 180.8	DB 15	Length 150462	Matches 510	Conserve/altive	0	Mismatches 292	Indels 124	Gaps 3											
Db	731	AGGAGGAAGGAGCAAGTACGCTTGACCCGATTTACGTGCACAAATTGATGAGACTTTGG	790	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27185,27995..28148))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (join124491..24817,24913..25266,25398..25633,25720..25832,27054..27185,27995..28147))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK062255"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor
Db	12351	AGGAAGTGAAGCAAGCAATCTGCACCTGATCTGCTTCACAGATGATGAACCTCTGG	12292	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	791	AAGAGTCACTCCGCGCTTATGTCCTGAGACTACTGGCTTACCGCTGCTGATGATTAACG	850	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	12291	AGGAGATTACACCTCGCTGATGATGAGACTTCTCCCTTCCTATGACACAGAGCAATC	12232	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	851	CTGCGAAAACCTAATGCTTTAAGCGGTGTGCGAATATTTTGTGCTGTTGAGAGAG	910	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	12231	ATAGAAGCCCAAGAGGAGCTTCAGAGGCGAGAAACATTTGTGAGACCTGGTGCACAG	12172	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	911	GTGAGACATCAGCTCTGTTGGGGGTTTGACCCGCTGAGAAAGTTTATGATGAGCGCTTT	970	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"	/note="contains full-length cDNA (s) : AK065148"	non-coding transcript probably inactive due to including stop codon(s) in CDS"	complement (join124787..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="contains EST (s) : AU031017(E60571), AU093466(E60571)"	contains full-length cDNA (s) : AK121372,AK062255"	/codon_start=1	/product="putative arginine/serine-rich splicing factor			
Db	12171	GAGGTATTGCTACCGTTGGAGAGAAATTTCTGTGTAACCTTCATGAACGAGGCTTTT	12112	complement (join124479..24817,24913..25266,25398..25633,25720..25832,27054..27159))	/gene="P0575F10.6-1"	/note="supported by full-length cDNA (s) : AK121372"	complement (24563..28147)	/gene="P0575F10.6-2"	complement (24563..28147)	/gene="P0575F10.6-2"											

*	33495	35494: contig of 3000 bp in length
*	35495	35594: gap of unknown length
*	35595	40051: contig of 4457 bp in length
*	40052	40151: gap of unknown length
*	40152	41379: contig of 1228 bp in length
*	41380	41479: gap of unknown length
*	41480	44517: contig of 3038 bp in length
*	44518	44617: gap of unknown length
*	44618	45564: contig of 947 bp in length
*	45565	45664: gap of unknown length
*	45665	49440: contig of 3776 bp in length
*	49441	49540: gap of unknown length
*	49541	52698: contig of 3158 bp in length
*	52699	52798: gap of unknown length
*	52799	55964: contig of 3166 bp in length
*	55965	56064: gap of unknown length
*	56065	59236: contig of 3172 bp in length
*	59237	59336: gap of unknown length
*	59337	70052: contig of 10716 bp in length
*	70053	70152: gap of unknown length
*	70153	84603: contig of 14451 bp in length
*	84604	84703: gap of unknown length
*	84704	86418: contig of 1715 bp in length
*	86419	86518: gap of unknown length
*	86519	88069: contig of 1551 bp in length
*	88070	88169: gap of unknown length
*	88170	89661: contig of 1492 bp in length
*	89662	89761: gap of unknown length
*	89762	91357: contig of 1596 bp in length
*	91358	91457: gap of unknown length
*	91458	132727: contig of 41270 bp in length
*	132728	132827: gap of unknown length
*	132828	137468: contig of 4641 bp in length
*	137469	137568: gap of unknown length
*	137569	163843: contig of 26275 bp in length
*	163844	163943: gap of unknown length
*	163944	169588: contig of 5915 bp in length
*	169589	169958: gap of unknown length
*	169959	176100: contig of 6142 bp in length
*	176101	176200: gap of unknown length
*	176201	183079: contig of 6879 bp in length
*	183080	183179: gap of unknown length
*	183180	184752: contig of 1573 bp in length.
FEATURES		Location/Qualifiers
source		1..184752
		/organism="Zea mays"
		/mol_type="genomic DNA"
		/strain="B73"
		/db_xref="taxon:4577"
		/clone="ZMMBC0216K08"
		2211..2310
gap		/estimated_length=unknown
gap		9942..10041
gap		/estimated_length=unknown
gap		16746..16845
gap		/estimated_length=unknown
gap		32395..32494
gap		/estimated_length=unknown
gap		35495..35594
gap		/estimated_length=unknown
gap		40052..40151
gap		/estimated_length=unknown
gap		41380..41479
gap		/estimated_length=unknown
gap		44518..44617
gap		/estimated_length=unknown
gap		45565..45664
gap		/estimated_length=unknown
gap		49441..49540
gap		/estimated_length=unknown
gap		52699..52798
gap		/estimated_length=unknown
gap		55965..56064

	gap	/estimated_length=unknown 59237..59336 /estimated_length=unknown
	gap	70053..70152 /estimated_length=unknown
	gap	84604..84703 /estimated_length=unknown
	gap	86419..86518 /estimated_length=unknown
	gap	88070..88169 /estimated_length=unknown
	gap	89662..89761 /estimated_length=unknown
	gap	91358..91457 /estimated_length=unknown
	gap	132728..132827 /estimated_length=unknown
	gap	137469..137568 /estimated_length=unknown
	gap	163844..163943 /estimated_length=unknown
	gap	169859..169958 /estimated_length=unknown
	gap	176101..176200 /estimated_length=unknown
	gap	183080..183179 /estimated_length=unknown
	ORIGIN	
	Query_Match	6.8%; Score 163.4; DB 14; Length 184752;
	Best Local Similarity	54.0%; Pred.No. 1.8e-32; Indels 135; Gaps 3
	MATCHES 506; Conservative	0; Mismatches 296;
OY	731 AGAGGAAGAACAATGACCTTGGACCAGATTTCGTCAAAATTTGATGACTTTGG	790
Dd	108936 AGAGGAATGTGCACACATCTTGACCGTAAGTCTGCCAAGATTGATAAATTGG	108999
OY	791 AAGAGTACTCCGCGTATATGCTTTGAGACTTACTTGGCCTTAGCCCTTGATGATTACG	850
Dd	108996 AGAGATTATCACCTGTGTGATTGGAGCTTCTGCTCTACTCATGAAAAATATA	109055
OY	851 CTGCGAAAAGCTAATATGTTTTAACCGGTGCGGAATTTTGTGCTGTGGAGAG	910
Dd	109056 AAAATTAACGGCAAAGAGCTGCAAGGTGCACAAAACATATTTGTGAGTGTGGACGG	109115
OY	911 GTGAGCATCAAGCTCTGTGTGGGGTTTGAACCCGTGAGAGTTTATGAAATGAGCCCTTT	970
Dd	109116 GTGGAATGTCTACTGTGAGAGAGATTTCCTGCGAGGCCCTTCAATGATGAGCCTTCT	109175
OY	971 TACGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCT-----	1011
Dd	109176 TGCAAGATCACTACAGCTGAGCAGGTTGTGCTACCAAGTTTATGACAACTACTGYTA	109235
OY	1012 -----	1011
Dd	109236 GTGTGTAGTCTGTCATTTTAATGACAGAGTTTATGAGCTTCTACTCATTTTTCTGT	109295
OY	1012 -----ACCCAAGCAATATTCACGAGAGTCAATTG	1042
Dd	109296 TTATATCTCAATGATTTCTTCTTAACACCGAATGCAATCCACTGATGATGTTTG	109355
OY	1043 AAGTTTACGAAGTTGCACTTGTCTTGTGGCTCAAGCTTTTATTTGTTAGAACCAACC	1102
Dd	109356 AGATCTATAGTGTGCACTTGCCCCACATTTGCTCAAGCAATTCAGATTAAGCCCACAAT	109415
OY	1103 TTTTACAGAGATCTGATAGCAATTTCCAGCAATTCACAGGCTTAAGTAATGCGTATGG	1162
Dd	109416 TCATCATATATGCGAGATTAACCTCTTGAAACCACTTCAGAAAGTTCAA-----	109461
OY	1163 AGATTCCTGCGATGTGTATGATACAGGAATTAATTGGAGATGAGCTTCGGCTTGAAAT	1222
Dd	109462 ----TATAGTTCCTCAATATCTTATGAAAAAGAGATGACCTTCGCAATTGAAA	109511

```

QY      1223 GGGAGCTGTGTGCACTGCTTATAGGCAAGTGTATGATCCGTATGTGTTGGCTTAG 1282
          |||||
Db      109512 GGGCACTGTCTGTTACTTGTGGGGATATTTGCAATGTCAGAAATGTGGCTTGGAAATTG 109571
QY      1283 ACGTGTGAGTTTCAATATATAGGATTCAGCTATTTGGAGTTTGTGTTGGAAATTCAA 1342
          |||||
Db      109572 ATATAGATTCCTGTCTATATAGGACCCCAAAATATCGAGTTTGTGTGTAATTAACCTTA 109631
QY      1343 ATGTG---GATGCAATGATGATCTCCCTGAGCTATGCAATTTGTTGGAAACCTGGTTGG 1399
          |||||
Db      109632 GCATGAAAGAAAGAAATGATCTTCTTCAGAGGCTGTGCAAGCTCTTGGAAACATGGCTTG 109691
QY      1400 CAGGGGTGTCTTCTTCTAGGTTCAAGACACCAAGATMAAAATTTAACTCGGGGACT 1459
          |||||
Db      109692 TCTCCGAGATATTCCTCCAGAGACAGATATCCGAGGAGACGAGTTCACTGGAGACT 109751
QY      1460 ACTATGATGATCTCTATGTTTGTGATTTGTTGAAAGAGTGAAGTTTCAAGGTTCTC 1519
          |||||
Db      109752 ATTACGACGACCAAAAGTTCTAAGCTTACGAAAGATGGAAGCGGTGTGCTTCCC 109811
QY      1520 CTTTACGCTGCTGCTGCAACTATGSCAAGATTTGAGC 1556
          |||||
Db      109812 ATTACGCTGACGCTGCCCTATAGCAAAACTTGTGC 109848

RESULT 12
AC160013 117818 bp DNA linear HTG 02-JUN-2005
LOCUS Medicago truncatula clone mth2-162b23, WORKING DRAFT SEQUENCE, 4
DEFINITION ordered pieces.
AC160013
VERSION AC160013.10 GI:66865008
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
            Burkarijota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
            Medicago.
REFERENCE 1 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-162b23
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2005) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE 3 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2005) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT On Jun 2, 2005 this sequence version replaced gi:66841540.
            ----- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

```

```

FEATURES
source
1..117818
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cd_xref="taxon:3860"
/clone="mth2-162b23"
/clone_1fb="Medicago truncatula BAC library H2"
1816..18915
/estimated_length=unknown
21098..21197
/estimated_length=unknown
35202..35301
/estimated_length=unknown

ORIGIN
Query Match 6.0%; Score 144.2; DB 14; Length 117818;
Best Local Similarity 68.3%; Pred. No. 2,9e-27;
Matches 200; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY      2114 AGGTTTGGATGGGCGCATGCTGAAGATTGGACTGACAGACAGCTGAACTGGCGCAGC 2173
          |||||
Db      1085 AGGTGTGGATCGGCGCAAAATTTGAAATATGATGATGACAGCTGAGATTTGCAAGC 1144
QY      2174 TTGGTGTGTTATGATATATATACACTGTGAAACATCTGTTGACAGTGTGACGTTCTCAG 2233
          |||||
Db      1145 TTGGTGTGATATATGATGATCACTCACTTGAAGATCTCAACATGACAGTGTGACCATATATCAG 1204
QY      2234 CAGATGAAACCCGTGCTGTGTGTGGAAGACAACTCTGGAAGAGTCTGTTCTATCTGATT 2293
          |||||
Db      1205 AGAATGGGGGGGGGTGAGTAGTAGGAACAACTCTAAGAGTCTACCACTCATCTGCTG 1264
QY      2294 TGGTTCATCCAGAAACAAATGCTACTGATGTGAGAACTACACAACTGATGAGAAATT 2353
          |||||
Db      1265 TTGGTATTCACAGCATGTCTACTTCCACAGCAGAACTCAACACAGATATGAAATGT 1324
QY      2354 TCTGTCCAGTCAAGGTGAGGTAATTCATGAAAGCTGTGTTGATCATATA 2406
          |||||
Db      1325 CTTTTCAGATTCAGGGTGAATAATATATGAAGAGCTGCTGTGATGTGTA 1377

RESULT 13
AC094893/c 221756 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-6B10, *** SEQUENCING IN PROGRESS ***
AC094893
VERSION AC094893.5 GI:30466744
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
            Burkarijota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 221756)
AUTHORS Muzny,D.,Marle., Metzker,M.,Lee., Adamson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,D.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Day-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Baccotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```


[illegible]

TITLE	JOURNAL
REFERENCE	AUTHORS
TITLE	JOURNAL
Unpublished 2 (bases 1 to 232974) Morley,K.C. Direct Submission Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 232974) Rat Genome Sequencing Consortium. Direct Submission Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On May 13, 2003 this sequence version replaced gi:24942402. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('a','contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	Stearns,M., Strong,R., Sutton,A., Stacey,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K., Valae,R., Vera,V., Villanena,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Morley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinsteck,G. and Gibbs,R.A. Direct Submission
Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information ----- Center project name: GLOG Center clone name: CH230-228E13 ----- Summary Statistics ----- Assembly program: Atlas 3.0; Consensus quality: 198057 bases at least Q40 Consensus quality: 202743 bases at least Q30 Consensus quality: 205785 bases at least Q20 Estimated insert size: 207912; sum-of-contigs estimation Quality coverage: 5x in Q20 bases; sum-of-contigs estimation	* NOTE: Estimated insert size may differ from sequence length * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html) . * NOTE: This is a 'working draft' sequence. It currently * consists of 4 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. 1 227401: contig of 227401 bp in length * 227402 227501: gap of unknown length * 227502 228773: contig of 1272 bp in length * 228774 228873: gap of unknown length * 228874 231053: contig of 2180 bp in length * 231054 231153: gap of unknown length * 231154 232974: contig of 1821 bp in length. Location/Qualifiers 1..232974 /organism="Rattus norvegicus" /mol_type="genomic DNA"

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2006, 13:05:25 / Search time 7918 Seconds
(without alignments)
4733.071 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGLSPQLCRLP.....YEVFWKSGMKITGVSVAS 801

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DRV=xlh
-Q=/abs/ABSSWEB.spool/US10600070/runat_21022006_112400_8653/app_query.fasta_1
-DB=EST -QFMT=fastlap -SUFFIX=2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abbs604
-USER=US10600070 @CGN_1_1_5315@runat_21022006_112400_8653 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hcc: *
5: gb_hcc: *
6: gb_est4: *
7: gb_est5: *
8: gb_est6: *
9: gb_est7: *
10: gb_gse1: *
11: gb_gse2: *
12: gb_gse3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1761.5	43.4	2307	10	CL965374
2	1529	37.6	919	4	CNS099YXH
3	1038	25.5	741	4	CNS099YXH
4	1038	25.5	741	4	CNS099YXH
5	1002.5	24.7	813	7	CO071968
6	970	23.9	897	6	CD573714
7	881	21.7	785	9	B2437564

8	794.5	19.6	769	7	CO075595	CO075595 GR_Ea36G
9	789.5	19.4	871	10	DU050213	DU050213 14781.T0
10	788	19.4	832	7	CO079829	CO079829 GR_Ea42O
11	784	19.3	561	1	AI998415	AI998415 701545606
12	776	19.1	534	5	BQ834167	BQ834167 ALB10128
13	774.5	19.1	703	7	CN918842	CN918842 030207ABP
14	767.5	18.9	741	8	CX543653	CX543653 UCRPT01_5
15	765	18.8	954	10	C3961431	C3961431 MBK428TR
16	738	18.2	898	8	DR830109	DR830109 ZM_BPD007
17	733.5	18.1	631	5	BU046755	BU046755 PP_LB0002
18	720.5	17.7	849	7	CO117047	CO117047 GR_EB019
19	718.5	17.7	624	8	CX195205	CX195205 13-E01936
20	718	17.7	660	2	BI268376	BI268376 NF119C111
21	716.5	17.6	697	8	CX173275	CX173275 BC01_69-11
22	694.5	17.1	728	8	CX674370	CX674370 UCRG10_6
23	671	16.5	611	8	CX117198	CX117198 B06_65-75
24	670	16.5	574	1	AW696905	AW696905 NF110C045
25	670	16.5	751	11	CR486014	CR486014 mch2-1550
26	652	16.0	820	8	DR972351	DR972351 ZM_BPD009
27	629	15.5	744	6	CD903230	CD903230 G356_109J
28	610.5	15.0	711	6	CF445760	CF445760 EST682105
29	609	15.0	646	3	BU464825	BU464825 B164825
30	605.5	14.9	593	7	CN904734	CN904734 010926ABD
31	604	14.9	922	7	CV264112	CV264112 MS02023.B
32	601.5	14.8	722	7	CO117046	CO117046 GR_EB019
33	596	14.7	664	8	DN939691	DN939691 6602_2 AE
34	590	14.5	819	7	CV241483	CV241483 MS02512.B
35	586	14.4	821	7	CK090561	CK090561 F013P64.3
36	573.5	14.1	723	7	CO079828	CO079828 GR_Ea42O
37	570	14.0	539	8	CX107243	CX107243 B113C23
38	563	13.9	866	6	CD573715	CD573715 UCRPT01_0
39	558	13.7	839	7	CK936162	CK936162 CGF100438
40	557.5	13.7	734	8	CX617940	CX617940 GABR1_36
41	350	13.5	746	8	CX617469	CX617469 UCRG10_6
42	548.5	13.5	521	7	CN914655	CN914655 030110ABN
43	548	13.4	768	8	CX543654	CX543654 UCRPT01_5
44	546	13.4	622	1	AV833644	AV833644 UCRPT01_5
45	544	13.4	415	3	BP785511	BP785511 BP785511

ALIGNMENTS

RESULT 1
LOCUS CL965374 2307 bp DNA linear GSS 21-SEP-2004
DEFINITION OBIFCC012338 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL965374
VERSION CL965374.1 GI:52385433
KEYWORDS
SOURCE
ORGANISM
GSS.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS 1 (bases 1 to 2307)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W., and Wang, J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL unpublished (2004)
COMMENT
Contact: Chen Chen
Department of Bioinformatics
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1..2307

FEATURES
source

ORIGIN

```

/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/notes="Oryza sativa exon trapped genomic sequences"

```

Alignment Scores:

Pred. No.:	2e-182	Length:	2307
Score:	1761.50	Matches:	390
Percent Similarity:	61.8%	Conservative:	119
Best Local Similarity:	47.4%	Mismatches:	213
Query Match:	43.4%	Indels:	101
DB:	10	Gaps:	17

US-10-600-070b-2 (1-801) x CL695374 (1-2307)

```

Oy 12 SerProPheGlnLeuCybArGLeuProAlaThrThrLybLeuArGArSerH1Asn 31
    ::::::::::::::::::::
Db 40 GCCCATTCGCCCTTCCCTCCCTCCGCCGCCGCCGCCGCCGCCGCCGCCGCCCT 99
Oy 32 ThrSerThrThrLeCyb--SerAlaSerLybTrpAlaAspArGLeuLeuSerAspPhe 50
    ::::::::::::::::::::
Db 100 CACCCCTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 159
Oy 51 AsnPhe-----ThrSerAspSerSerSerSerPheAlaThrAlaThr 65
    ::::::::::::::::::::
Db 160 CACCTCTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 219
Oy 66 ThrThraThrLeuValSerLeuProProSerL1eAspArGProGluArGHisValPro 85
    ::::::::::::::::::::
Db 220 GCGCCCTCCGCCCTCCCTTCGTCGCGCTTCGCCGAGCGCGCGCGCGCGCGCGCTCC 279
Oy 86 IleProIleAspPheTyrgln-----ValLeuGlyAlaGln 97
    ::::::::::::::::::::
Db 280 CTCACAGTCGATTTCTCAAGAGTGAGGCGCCCGTGAATCTGCTGCGGTTCTAGGGCG 339
Oy 98 ThrHisPheLeuThrAspGlyIleArGArGAlaPheGluAlaArGValSerLybProPro 117
    ::::::::::::::::::::
Db 340 CCACATTTCTCTGCGCGATGCGATCAAGAGGCGCTTCAGAGCGAGATAGCCAGCCACCG 399
Oy 118 GlnPheGlyPheSerAspAspAlaLeuIleSerArGArgGlnIleLeuGlnAlaIaCyb 137
    ::::::::::::::::::::
Db 400 CAGTATGGCTACAGCAAGCATGCTCTTGTGTCGTGTCGACAAATGCGAGATGGCCAT 459
Oy 138 GlnThrLeuSerAsnProArGSerArGArgGlnTyrglnGlnGlyLeuLeuAspAspGln 157
    ::::::::::::::::::::
Db 460 GACACTCTCATGAACCAAGAACTCCCGCACTCAGTATGATGTCGCGCTTCTGAGAACCGT 519
Oy 158 GlnAlaThrValIleThrAspValProTrpAspLybValProGlyAlaLeuCybValLeu 177
    ::::::::::::::::::::
Db 520 GAAGAAAGCTCTCACCATGATATTGCTTGGAGACAG----- 555
Oy 178 GlnGlnGlyGlyGlnThrGlnIleValLeuArGValGlyGlnAlaLeuLeuLybGluArG 197
    ::::::::::::::::::::
Db 556 ---GAGGCTGGGAGGAGCATCTGCTGCTTGTAACTGGAGAACAGTTGCTTGGAGTCGG 612
Oy 198 LeuProLybSerPheLybGlnAspValValLeuValIleAlaLeuAlaPheLeuAspVal 217
    ::::::::::::::::::::
Db 613 CCACCCAAAGCGCTTCAAGAGCAGACGTGTCTAGCGATGCTGCTGTATGTGAGACTA 672
Oy 218 SerArGAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrglnPheValGlu 237
    ::::::::::::::::::::
Db 673 TCAGGAGATCCTATGGCAGAACCCCTCCAGATGTAATGGCTGCTCGAGAGGTCTGAG 732
Oy 238 GlnAlaLeuLybLeuGlnGlnGlnGlyValaSerSerLeuAlaProAspLeuArGAla 257
    ::::::::::::::::::::
Db 733 AGGCTCTCAAGCTCTTTCAGAGAGATGAGACAAAGCAATCTCGACTGATCTCTTCA 792
Oy 258 GlnIleAspGlnThrLeuGlnGlnGlnIleThrProArGTyrglnValLeuGlnLeuGlyLeu 277
    ::::::::::::::::::::
Db 793 CAGATTGATGAATCTCGAGAGATTAACACTCGCTGTGTAATGGAGCTTCTCTCCCTT 852

```

```

Oy 278 ProLeuGlyAspAspTyrglnAlaLybArGLeuAenGlyLeuSerGlyValIleArGAsnIle 297
    ::::::::::::::::::::
Db 853 CCAATTGACACAGAGCATCATAGAAGCCGCCAAGAGGGCTTCAAGGCTCGAGAAACATT 912
Oy 298 LeuTrpSerValGlyGlyGlyValaSerAlaLeuValGlyGlyLeuThrArGLeuLyb 317
    ::::::::::::::::::::
Db 913 TTGTGAGCGGTGGCAGAGAGATGTGTACCGGTGGAGAGATTTTCTGTGAGGCC 972
Oy 318 PheMetAsnGlnAlaPheLeuArGMetThrAlaAlaGlnGlnValaAspLeuPheValAla 337
    ::::::::::::::::::::
Db 973 TTCATGACAGAGCTTTTGTAGAGATGACATCAATGMAACAGATGATTTCTTTCAAA 1032
Oy 338 ThrProSerAsnIleProAlaGlnSerPheGlnValTyrglnValaLeuAlaLeuVal 357
    ::::::::::::::::::::
Db 1033 ACACCGAATAGCATTCCTCTCAATGTGTAATTAACATGACCTTGACATGTC 1092
Oy 358 AlaGlnAlaPheIleGlyLybProH1eLeuGlnAspAlaAspLybGlnPheGln 377
    ::::::::::::::::::::
Db 1093 GCTCAAGCAATTATTAAGTAAAGCCCAATTATCATGATGCGGATGATCTTTTGA 1152
Oy 378 GlnLeuGlnAlaLybValMetAlaMetGlnIleProAlaMetLeuTyrglnArG 397
    ::::::::::::::::::::
Db 1153 CAACTCAGAAAGTTCAACATAGTTCT-----CATTAAGCTTATGATATAT--- 1197
Oy 398 AsnAsnTrpGlnIleAspPheGlyLeuGlnArGlyLeuCybAlaLeuLeuIleGlyLyb 417
    ::::::::::::::::::::
Db 1198 -----GAGATGACCTTGCAATGCAAAAGGCATTCTGCTCATGTGCTAGTCGAGAT 1248
Oy 418 ValaAspGlnCybArGMetTrpLeuGlyLeuAAspSerGlnaAspSerGlnTyrglnArGAsnPro 437
    ::::::::::::::::::::
Db 1249 GTTAGCAAGTGCAGAAATGTGCTGGAATGTAATGATCTTACACATACAGAGACCC 1308
Oy 438 AlaIleValaGlnPheValLeuGlnAenSerAsn---ArgAspAspAsnAspLeuPro 456
    ::::::::::::::::::::
Db 1309 AAAATTCTAGAGTTATTTGTGACCACTAGACATGAGAGAGATATGCTTCTTCCA 1368
Oy 457 GlyLeuCybLybLeuLeuGlnIleThrTrpLeuAlaGlyValaPheProArGpheaAsp 476
    ::::::::::::::::::::
Db 1369 GGGCTGTGAAAGCTTTTGGAGACTTGCTTATCTTGAAGTTTCTTGTGAGACAGAT 1428
Oy 477 ThrLybAspLybPheLybLeuGlyAlaAspTyrglnArGAspProMetValaLeuSerTyrg 496
    ::::::::::::::::::::
Db 1429 ACTCGGGGCATGACGATTCAGACTTGAGATTAATGATTCAGAGATTCAGAAAGTTTAAGCTAC 1488
Oy 497 LeuGlnArGValaGlnValaGlnGlySerProLeuAlaAlaIleAlaThrMetAlaArG 516
    ::::::::::::::::::::
Db 1489 CTGAAAGAGATGAGAGGTGTGTGCTTCTCATTTGGCTGCTGCTGCTGCTATTCGAAAA 1548
Oy 517 IleGlyAlaGln-----HisVallybAlaSerAlaMetGlnAlaLeu 530
    ::::::::::::::::::::
Db 1549 CTGTGTCTCAAGCTACAGCTGCACTTGTAAGTGAATCAATGCTATTCAGAGGTTTC 1608
Oy 531 GlnLybValPhePro---SerArGTyrglnThrAspArGAsnSerLybGlnProLybAspVal 549
    ::::::::::::::::::::
Db 1609 AACAAAGTTTTTCAATGATAGAAACAGTAAAGAGGTGAGCCATGAGAAATACT----- 1662
Oy 550 GlnGlnThrValaPheSerValaAspProValaGlyAsnAsnValaGlyArGAspGlyGlnPro 569
    ::::::::::::::::::::
Db 1663 ----- 1701
Oy 570 GlyValPheIleAlaGlnAlaValaArgProSerGlnaAsnPheGlnThrAsnAspTyrgAla 589
    ::::::::::::::::::::
Db 1675 GGGGATATCTT-----GAAATTTTGAACAG----- 1701
Oy 590 IleArGAlaGlyValSerGlnSerSerValaAspGlnThrThyValaGlnMetSerValAla 609
    ::::::::::::::::::::
Db 1702 -----GAAATGCACTGCT 1716
Oy 610 AspMetLeuLybGlnaAspValLybIleLeuAlaIaGlyValaIleGlyLeuIle 629
    ::::::::::::::::::::
Db 1717 CATGATTCGAGAAATGCGCTTGAAATTAATCTCTGCTGCGCATGTTTGCACGTGTC 1776
Oy 630 SerLeuPheSerGlnTyrglnPheLeuLybSerSerSerSerPheGlnArGlybAspMet 649

```

```

Db      1777 GCAGTAAATGGGGCCCAATAT-----TTGCCCTCGTAAGAGGCC 1815
      ::::|||||
Qy      650 ValSermetetGuserAep-----ValAlaThrileGlySerValArgAlaAsp 666
      ::::|||||
Db      1816 CTTTCCTGCTATTAGAGTGAAGCATGATCTGTGGCAGTCTGTAATGATGCTGACTACT 1875
      ::::|||||
Qy      667 AspSerGluAlaLeu-----ProArgMeCAAPAlaArgThrAla 679
      ::::|||||
Db      1876 GATGATCTCTGCACCTAGTGAAGATCCAGTACATATCTTCAATAGATGCGAAGCTGGCA 1935
      ::::|||||
Qy      680 GluAsnIleValSerIleSTrGInLysIleLysSerLeuAlaPheGlyProAspPheArg 699
      ::::|||||
Db      1936 GAAGATATGTTGTCAGAGTGGCAGAGTATCAATCTAAGGCTTGAGGACCGAAGCATTCG 1995
      ::::|||||
Qy      700 IlleGluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTrpThrAspArgAla 719
      ::::|||||
Db      1996 GTTGCATCATGTCAGAGGTTCTTGTATGGCAACATGCTTAAAGGTGTGACTGACCGAGCA 2055
      ::::|||||
Qy      720 AlAGluThrAlaGluLeuGlyLeuValTrpAspTrpThrLeuLeuLysSerValAsp 739
      ::::|||||
Db      2056 GCGGAGATTGAGGCTCATGGGTGTTCTGGAGATACACTATCCGATGACGATTGAT 2115
      ::::|||||
Qy      740 SerValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSer 759
      ::::|||||
Db      2116 AGCATCACTATCTCTCCATAGTGTGCGAGGCGACTGTGAGGCTTACGATTGATGAGGCA 2175
      ::::|||||
Qy      760 AlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThr 779
      ::::|||||
Db      2176 GGGCAACTTACTGATGTTACTGAGCCAGAAACATATCTCATATGACACAAATACACT 2235
      ::::|||||
Qy      780 ThrArgTrpGluValPheTrpSerIle-----SerGlyTrpLysIleThrGluGluSerVal 798
      ::::|||||
Db      2236 ACCGGATGATGAGAGGCTTCTCCAGCTAGAGAGGTGAGATACGAGGAGGAGCTC 2295
      ::::|||||
Qy      799 IeuAlaSer 801
      ::::|||||
Db      2296 CTCAGATCG 2304
      ::::|||||

RESULT 2
CNS09YJH
LOCUS
DEFINITION
CNS09YJH 919 bp mRNA linear HTC 04-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTJS892C08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX841670.1 GI:42406830
VERSION
BX841670.1
KEYWORDS
HTC; GSTR cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 919)
REFERENCE
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Queirer,F., Scarcelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
JOURNAL
2 (bases 1 to 919)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLES
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

```

```

FEATURES
source
location/Qualifiers
1..919
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTJS892C08"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..919
/gene="At5g42480"
/origin="At5g42480"

Alignment Scores:
Pred. No.: 2,1e-157 Length: 919
Score: 1529.00 Matches: 301
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 37.6% Indels: 0
DB: 4 Gaps: 0

US-10-600-070b-2 (1-801) x CNS09YJH (1-919)
Qy      1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGluLeuCyAspArgLeuPro 20
      :|||
Db      13 ATGGAAGCTGTGAGTACGTCGGCATTTGCTCTCCCATTTCAATATATCCGATTACCA 72
      :|||
Qy      21 ProAlaThrThrIleValLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
      :|||
Db      73 CCGGCGACGACAAAGCTCCGACATAGCCACAAACCTCTCAACTATCTCTCCGACG 132
      :|||
Qy      41 LysThrAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
      :|||
Db      133 AAATGGGCGACGCTCTCTCCGACTTAATTTACCTCCGATTCCTCTCTCTCTCC 192
      :|||
Qy      61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
      :|||
Db      193 TTGGCCACCGCCACACACACGACCTCTGCTCTCCGACACATATATGATGATGCC 252
      :|||
Qy      81 GluArgHisValProIleProIleAspPheTrpGluValLeuGlyAlaGlnThrHisPhe 100
      :|||
Db      253 GAAGCGACGTCGCCCATCCCATTTGATTTCTACCAAGTATTAAGAGCTCAAAACATTTTC 312
      :|||
Qy      101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
      :|||
Db      313 TTAAACGATGAATTCAGAAAGATTCGAAGGTGAGGTTTCGAAACCGCGCAATTCGGT 372
      :|||
Qy      121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
      :|||
Db      373 TTACGACGACACGCTTTTACACCGGACGACAGATTTCTTAAGCTGCTTCCGAAACTCTG 432
      :|||
Qy      141 SerAspProArgSerArgArgArgIleValSerGluGlyLeuLeuAspAspGluAlaThr 160
      :|||
Db      433 TCTAATCTCGGCTGTAAGAGAGTCAATGAGAGGCTCTCTGATGATGAAGAGCTACA 492
      :|||
Qy      161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuLysValLeuGlnGluGly 180
      :|||
Db      493 GTATCACTGATGTTCTTGGATGAAGTTCTCGGTGCTCTGCTGATGATGAAGAGT 552
      :|||
Qy      181 GlyIleThrGluIleValLeuArgValGlyAlaLeuLeuLysGluArgLeuProLys 200
      :|||
Db      553 GGTGAGACTGAGATAGTCTTCCGGTTGGTGGAGCTCTGTTAAGAGAGGTTCCCTAAG 612
      :|||
Qy      201 SerPheLysGluAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
      :|||
Db      613 TCGTTTAAAGCAAGATGTTGTTAGTTATGCGCGCTTCTCGATGTCGAGAGGAT 672
      :|||

```

QY	221	AlaMet1aLaAaSPProPProAaSPhe11eThrGlyTYrGluPheValGluGlu1A1eU	240
Db	673	GCTATGCAATTGGATTCACCTGATTTTAACTGTTATGACTTTGTTAGAGAAAGCTTGG	732
QY	241	LysLeuLeuGingJugJug1ValAspSerLeuAlaProAaPLeuArgAlaGln1LeaP	260
Db	733	AAGCTTTTACAGAGAGAGCAAGTACGCTTGACCGGATTTTACGTCACCAATTCAT	792
QY	261	GluThrLeuGingJug11eThrProArgTYrValLeuGluLeuGluGly1LeuProLeuGly	280
Db	793	GAGACTTTGGAGAGAGATCACTCCGGTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT	852
QY	281	AaPaAPTYrAlaAla1ArgLeuAsnGlyLeuSerGlyValArgAsn1LeuTrpSer	300
Db	853	GATGATTTACGCTGCGAAGAAAGACTAAATGCTTTAAGCGGTGTCGGAATATTTTGCGCT	912
QY	301	ValGly	302
Db	913	GTTGGA	918
RESULT 3			
CNSOBYNM			
LOCUS			
DEFINITION	CNSOBYNM	741 bp	mRNA linear HTC 06-FEB-2004
	Arabidopsis thaliana Full-length cDNA Complete sequence from clone		
	GSLSIR592A05 of Siliques of strain col-0 of Arabidopsis thaliana		
	(thale cress).		
ACCESSION	BM333489.1	GI:42455179	
VERSION	BM333489.1	GI:42455179	
KEYWORDS	HTC; GSLS_CDNA.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
	Rosidae; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 741)		
AUTHORS	Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,		
	Menard,M., Crnaud,C., Querlet,F., Scarpelli,C., Schachter,V.,		
	Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.		
	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:		
	A Combined Approach to Evaluate and Improve Arabidopsis Genome		
	Annotation		
	Unpublished		
	2 (bases 1 to 741)		
JOURNAL	Genoscope.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (18-NOV-2003) Genoscope - Centre National de Sequenage :		
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
COMMENT	The sequences are based on single pass reads.		
	Life Technologies (a division of Invitrogen) members carried out		
	full-length libraries construction : Temple G.		
	Genoscope members carried out sequencing and annotation : Castelli		
	V., Aury J.M., Jallion O., Wincker P., Menard M., Crnaud C.,		
	Schachter V., Weissenbach J., Salanoubat M.		
	URGV INRA : Clepet C., Caboche M.		
	Annotation is based on the June 2003 version of the Arabidopsis		
	genome released by MIPS (Munich Information center for Protein		
	Sequences). 5 prime and 3 prime are assembled with Phrap.		
	http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full		
	length		
	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.		
FEATURES			
source	1..741		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="GSLSIR592A05"		
	/tissue_type="Siliques"		
	/ecotype="Col-0"		
	/plasmid="pCMVSPORT_6"		
	1..741		
gene	/gene="At5g42460"		

ORIGIN

Alignment Scores:	2.37e-103	Length:	741
Pred. No.:	1038.00	Matches:	210
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	25.5%	Indels:	0
Query Match:	4	Gaps:	0
DB:			
US-10-600-070B-2 (1-801) x CNS091NM (1-741)			
QY	592	AlaGlyValSerGluSerSerValAspGluThrThrValGluMetSerValAlaAspMet	611
Db	1	GCTGGGGCTTCAGAGATAGCGTTGAGAACTACTCTTGAATNGTCGGTTGCGATATG	60
QY	612	LeuLysGluAlaSerValLysLeuAlaAlaGlyValAlaIleGlyLeuLysLeu	631
Db	61	TTAAAGAGGCAAGTGTGAAGATCTCTAGCTGCTGTGTGGCAATGTGACGATTCCTG	120
QY	632	PheSerGlnLysIleThrPheLeuLysSerSerSerPheGlnArgLysAspMetValSer	651
Db	121	TTTCAGCCAGAAAGTATTTCTTAAAGACGCTCACTTTTCAACCCAGAGATATGCTTCT	180
QY	652	SerMetGluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeu	671
Db	181	TCATGGAAATGTATGTCGCTACCATATGGCTCACTCAAGCTGACATTCAGAAAGCTT	240
QY	672	ProArgMetAspAlaArgThrAlaGluAsnIleValSerLysIleThrGlnLysLeu	691
Db	241	CCGAGATAGATGCTGACGATCGACAGAAATATAGTATCCAAATGGCAGAAAGATTAGCT	300
QY	692	LeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValIleAspGlyArgMet	711
Db	301	CTGGCTTTTGGGCGCTGATCACCGCATGAAATGTTCACAGAGGTTTGGATGGCGAATG	360
QY	712	LeuLysIleThrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValIleAspTyr	731
Db	361	CTGAAGATTTGGACTGACAGACGACGCTGAACTGCCAGCTGGGTGTTATGATAT	420
QY	732	ThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArgAlaLeu	751
Db	421	ACACTGTGGAACATATCTGTTGACAGGTGTGCACTTCAGACAGATGAAACCGGCTCTG	480
QY	752	ValGluAlaThrLeuGluGlnSerAlaCysLeuSerAspLeuValHisProGluAsnAsn	771
Db	481	GTCGAGAGCAACTCTGGAGAGCTGCTGTCTTACTATTTGCTTCAATCCAGAAACAT	540
QY	772	AlaThrAspValAlaGlyThrIleThrThrArgTyrGlnValPheTyrSerLysSerGlyTyr	791
Db	541	GCTACTGATGTGCAGAACTTACACAAAGATACAAAGATTTCCTGCTCAAGTCAAGGTGG	600
QY	792	LysIleThrGlnGlySerValLeuAlaSer	801
Db	601	AAATCACTGAAGCTCTGTTCTTGATCA	630
RESULT 4			
CNS0925S	741 bp	mRNA	linear
LOCUS	CNS0925S	Complete sequence from clone	
DEFINITION	GS1TSL1312E09 of Siliques of etrain col-0 of Arabidopsis thaliana		
	(thale cress).		
ACCESSION	BX833051		
VERSION	BX833051.1	GI:42455503	
KEYWORDS	HTC; GS1T cDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
AUTHORS	Castelli,V., Aury,J.M., Jallion O., Winkler,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,		

Db		260	GCAATGCTGGGGGATTACAGTAAACTTTTCAGTAATGAAGCCCTTTCTACGTATGCACA	319
Oy		328	AlaAalagluglnValAspLeuPheValAlaThrProSerAsnIleProAlaGluSerPhe	347
Db		320	TCAGCTGACGAGGTAACTATATTTCTGCGCACACCAAAADAGATCCGACGTGAACCTTT	379
Oy		348	GluValIYrGIuValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyIylasPheHis	367
Db		380	GAAAGCTTAGAGTGCGCATTGCCCTTGCCCAAAGCTTGTGGGTAGAACCACTCAT	439
Oy		368	LeuengInaPaAlaAspLysGlnPheGlngInuengIngnlAlaLysValMetAlaMet	387
Db		440	CTCATTCAGATNGCGATPAACATGTTCAAGCACCTTCAGCAAGATPAAGTACCAGCTTTA	499
Oy		388	GluIleProAlaMetLeuTYrAspThrArgAsnAntTrpGluIleaspPheGlyLeuGlu	407
Db		500	CGGACCTTGCGCTCATATCTATATTCCTCGAAAAAACATGAGATGCAATTCGACATTGAA	559
Oy		408	ArgGlyLeuCYsalAlaLeuLeuIleGlyLysValAspGluCYbarGmetTryLeuGlyLeu	427
Db		560	AGGGGCTCATGTCTCACCTCTGTAGGTAGAAGTCATGATGATGTGCTTATGTTGGGCTTA	619
Oy		428	AAsperGIuaPsSerGlnTYrArgAsnProAlaIleValGluPheValLeuGluAsnSer	447
Db		620	GATAGTGATPAAGTCCCACATACAGAAATCCAGCCATTTGATGATTTGTTGGAAAACCTCA	679
Oy		448	AsnArgAspAspAsnAspAspLeuProGlyLeuCYslsbleuGluThrTryLeuAla	467
Db		680	AAGCAAGCTAATGACAAATGATCTTCCTGCTCTCTTAAGCTTTAGACATAGCTGGCA	739
Oy		468	GlyValValPheProArgPheArgAspThrLysAspLysPheLysLeuGlyAspTYr	487
Db		740	GAGGTGTTTTCCCTAGATTCAGGGACACCAAGTGATATACGGTTCAACTGGATATAC	799
Oy		488	TyrAspAPPrometValLeuSerTYrIeuGluArgValGluValAlaGlnGlySerPro	507
Db		800	TATAGTATCCCTACTGTCCTTCGATATTTGAAACACGTGAGGAGGAATGATGATCGCCC	859
Oy		508	LeuAlaAlaAlaAlaThrMetAlaArgIleGlyAla	519
Db		860	TTAGCTGCAGACGCTGCCATGATTAGGCTTGAGCT	895
RESULT 7				
LOCUS	B2437564/c			
DEFINITION	BONRN72R BO_1.6_2_KB_cot Brassica oleracea genomic clone BONRN72,			
ACCESSION	B2437564			
VERSION	B2437564.1			
KEYWORDS	GSS.			
SOURCE	Brassica oleracea			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosoids; eurosids II; Brassicales; Brassicaceae; Brassica.			
REFERENCE	Aylee, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Uteback, T.R., Wortman, J.R., White, O.R. and Town, C.D.			
AUTHORS	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)			
TITLE	15805490			
JOURNAL	PUBMED			
COMMENT	Other GSSs: BONRN72RF Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TR Class: sheared ends.			

FEATURES

Source

```

location/Qualifiers
1..785
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="Tol000DH3"
/db_xref="taxon:3712"
/clone="BONNNT72"
/clone_1lb="BO_1.6.2_KB_tot"
/notes="vector: pHS1; site_1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHS1 using BstXI linkers"

```

ORIGIN

Alignment Scores:

Pred. No.:	5.59e-86	length:	769
Score:	881.00	Matches:	1951
Percent Similarity:	83.5%	Conservative:	6
Best local Similarity:	80.9%	Mismatches:	4
Query Match:	21.7%	Indels:	35
DB:	9	Gaps:	2

US-10-600-070B-2 (1-801) X BZ437564 (1-785)

[illegible]

DEFINITION GR_Ea36G17.r GR_Ea Gossypium raimondii cDNA clone GR_Ea36G17.3', mRNA sequence.

ACCESSION CO075595

VERSION CO075595.1 GI:48745076

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid 1; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 769)

AUTHORS Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C., and Wang, R. A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 36 Row: G Column: 17.
Location/Qualifiers

FEATURES

source

1..769

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ea36G17"

/cruise_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH108"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1. Site 1: Not; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into Not-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.:	1,92e-76	Length:	769
Score:	794.50	Matches:	157
Percent Similarity:	79.9%	Conservative:	34
Best Local Similarity:	65.7%	Mismatches:	41
Query Match:	19.6%	Indels:	7
DB:	7	Gaps:	4

US-10-600-070b-2 (1-801) x CO075595 (1-769)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProphGlnLeuCyArgLeu--- 19

DB 54 ATGAAATCTTTGAGGACATTAGCATGCGGTGTCACCTCTACACTTACTCTTTATTC 113

QY 20 ---ProProAlaThrThrIlyLeuArgArgSerHisAnThrSerThrThrIleCysSer 38

DB 114 CACCTCTCTAAACCTCCAGACTCCAC-----CACGGTCAACTACCGTCGTCTCC 167

QY 39 AlSerIyGTrpAlaAspArgLeuLeuSerAspPheAnPheThrSerAspSerSerSer 58

DB 168 GCCAGTAAATGGCGCGAGCGCTCTGCTGACTTCCAAATTTCTCCCTGCCCGGATAT 227

QY 59 SerSerPheAlaThrAlaThrThrAlaThrIleu---ValSerLeuProPheSerIle 77

DB 228 TCCGTCTCTCTCTCTCTCAACGCACTTTCTCTCCCTTACCCCTTACCCCTTCTC 287

QY 78 -----AspArgProGluArgHisValProIleProIleAspPheTyGlnValLeuGly 95

DB 288 GCCCTTCTCTCCCGCAACGCAAGCTTCCATTCCTTGAATTTCTCAAGGTTTGA 347

QY 96 AlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIys 115

DB 348 GCCGAGACTATTTCTTAGTGATGATGATCAGAGAGCTATGAAAGCAAGGTTTGA 407

QY 116 ProProGlnPheGlyPheSerAspAlaLeuIleSerArgGlnIleLeuGlnAla 135

DB 408 CCCCTCAATATGAGTTCAGTCAAGACACCATATTAAGCGAAGACAGATTCTTAGCT 467

QY 136 AlaCysGlnThrIleuSerAnPProAspSerArgArgGlnTyArgGlnIleuLeuAsp 155

DB 468 GCCTTGAAACCCATCTACCTGAGCTTACGCAAAATTAACAACAGGCTTGTGCAG 527

QY 156 AspGluGluAlaThrValIleThrAspValProTrpAspIlyValProGlyAlaLeuCy 175

DB 528 GATGACCGGACACTTATCATCTCACTCAAGTCCCTGGGACAGAGTCTTGAGCTGTC 587

QY 176 ValLeuGlnGlnGlyIleGlnThrGlnIleValLeuArgValGlyGluAlaLeuIys 195

DB 588 GTTTCAGAGAACTGCGGGAGACTGAAGTGTCTTCAATTTGGGAGAGTTTGTCAG 647

QY 196 GluArgLeuProIlySerPheIyGlnAspValValLeuValMetAlaLeuAlaPheLeu 215

DB 648 GAGAGGCTGCGCCCAAGCATTTAAACAGATGTCTGTCGCAATGCGCCCTTATGTG 707

QY 216 AspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyArgIu 234

DB 708 GACTTGCGAGGAGATGCTATGCTTGTGATCTTCAGATTTCATTGTGTGTGAG 764

RESULT 9

DU050213/c 871 bp DNA linear GSS 12-AUG-2005

LOCUS DU050213

DEFINITION 144781 Tomato HindIII BAC library Lycopersicon esculentum genomic clone LE HBa0013B21 5, genomic survey sequence.

ACCESSION DU050213

VERSION DU050213.1 GI:72480609

KEYWORDS GSS.

SOURCE Lycopersicon esculentum (Solanum Lycopersicum)

ORGANISM Lycopersicon esculentum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 871)

AUTHORS Mueller, L.A., Bueis, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J., Van Eck, J. and Stack, S.

TITLE BAC end sequencing from three Solanum Lycopersicon libraries

JOURNAL Unpublished (2005)

COMMENT Other GSSs: 144780

Contact: Lukas Mueller

Tanksley Lab, Dept. of Plant Breeding

Cornell University

251 Emerson Hall, Ithaca, NY 14853, USA

Tel: 607-255-6557

Fax: 607-255-6683

Email: sgn-feedback@sgn.cornell.edu

Insert length: 76680 Std Error: 0.00

Plate: 13 Row: B Column: 21

Seq primer: T7

Class: BAC ends

High quality sequence start: 43

High quality sequence stop: 789.

Location/Qualifiers

source

1..871

/organism="Lycopersicon esculentum"

/mol_type="genomic DNA"

/cultivar="Heinz 1706"

/db_xref="taxon:4081"

/clone="LE HBa0013B21"

/lab_host="E. coli"

/clone_lib="Tomato HindIII BAC library"

/note="Vector: pBelobAC11; Site_1: HindIII"

ORIGIN

Alignment Scores:

Pred. No.:	8,35e-76	Length:	871
Score:	789.50	Matches:	171
Percent Similarity:	70.1%	Conservative:	17

Best Local Similarity: 63.8% Mismatches: 40
 Query Match: 19.4% Indels: 40
 DB: 10 Gaps: 4
 US-10-600-070b-2 (1-801) x DU050213 (1-871)

```

Qy 234 GluPheVal-----GluGluAlaLeuYsLeu---LeuGln 244
    |||||
Db 843 GAATTTCTAGCTTACAGTGAAGCAATATATAGAGAACTTCTTCCCTTTTCCAG 784

Qy 245 GluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGlu 264
    |||||
Db 783 GAAGAAAGTCCAAATATCTTGCCCTTGATCTGCATCTCAGATCCAGATAGATGACATTAGAA 724

Qy 265 GluIleThrProArgGlyrValLeuGluLeuGlyLeuProLeuGlyAspAspTyrAla 284
    |||||
Db 723 GAATTAATCCAGCTGACGTAAGTGAACCTTCTTCTTCTTGATGATATACGA 664

Qy 285 AlaIleArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSerValGlyGly 304
    |||||
Db 663 ATGAAAAGAGTACAGGCTCTTCAAGTGTGCGCAACTTTGTGGCTGTGAGAGAGA 604

Qy 305 GlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysAspMetAsnGluAlaPheLeu 324
    |||||
Db 603 GGAGCAGCTCAATTTTCAGGGGGGTTCAACAGAAATTTTCAATGATGAGGCTTCTTA 544

Qy 325 ArgMetThrAlaAlaGluGln----- 331
    |||||
Db 543 CGATGACAGCTGTGAGCAGGTGGCTTAATTCAATTATTTTGGTTCTTTTGT 484

Qy 332 -----ValAspLeuPh 335
    |||||
Db 483 AACTGTTCTTTCACCTCTCTCTCAATTTCTCTTCTTTTAAAGGTGACCTCTT 424

Qy 335 eValAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAl 355
    |||||
Db 423 CGTCGACAGCCCAAGTAACTTCTGTCAGAAAGCTTTGAGGTTACCGGGTGGGCTTGC 364

Qy 355 AlaLeuValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysG 375
    |||||
Db 363 ACTGTGTGCTCAAGCTTTTGTGGAAAAAAACCTCATCTCATTCACATCTGATTAACCT 304

Qy 375 nPheGlnGlnLeuGlnAlaIleValysValIleAlaMetGluIleProAlaMetLeuTyrAs 395
    |||||
Db 303 TTTTCAGCAGCTTCAAGCAGACCAAGATTACAGCTTACGCACTTCTGTCTGATACAC 244

Qy 395 pThrArgAsnAsnTrpGluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuI 415
    |||||
Db 243 TGTTAAGAAACCGTGAATATAGCTTCGCTTGGAGAGGGGCTTTGTTCACTGCTTGT 184

Qy 415 eGlyLysValAlaAspGluCyAlaArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyr 435
    |||||
Db 183 TGAAGAAAGTGAAGTGAAGTGTGCTTGTGGGCTTGAAGAGTGAAGAGCTCCCTTAG 124

Qy 435 GAspProAlaIleValIleGluPheValLeuGluAsnSerAsnArgAspAsp---AsnAsp 454
    |||||
Db 123 AGATTCAGAGTATATGACTTTTGTGTCAGAACTCAAGAGACGACAAAGAAATGATCT 64

Qy 454 pLeuProGlyLeuCyAlaLeu 461
    |||||
Db 63 GCTCCTGACCTATGTAGCTT 42

RESULT 10
COO79829 832 bp mRNA linear EST 15-JUN-2004
LOCUS GR_Ea42007.r GR_Ea Goseypium raimondii cDNA clone GR_Ea42007 3',
DEFINITION mRNA sequence.
ACCESSION COO79829
VERSION COO79829.1 GI:48749310
KEYWORDS EST.
SOURCE Goseypium raimondii
ORGANISM Goseypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Goseypium.
 1 (bases 1 to 832)
 Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
 Wing,R.A.
 Global assembly of Cotton ESTs
 Unpublished (2004)
 CONTACT: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 Plate: 42 row: 0 column: 07.
 Location/Qualifiers
 1..832
 /organism="Goseypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Ea42007"
 /tissue_type="whole seedlings"
 /dev_stage="first true leaves"
 /lab_host="DH10B"
 /clone_lib="GR_Ba"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV; Library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Colonies
 plated/picked by AGI. More glycerol clones held in -80."

Alignment Scores:
 Pred. No.: 1.14e-75 Length: 832
 Score: 788.00 Matches: 172
 Percent Similarity: 67.6% Conservative: 30
 Best Local Similarity: 57.5% Mismatches: 61
 Query Match: 19.4% Indels: 36
 DB: 7 Gaps: 5

```

US-10-600-070b-2 (1-801) x COO79829 (1-832)

Qy 343 ProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaPheIle 362
    |||||
Db 3 CTTGCAAGAAAGTTTCAAGTTTATGTGTGACCTTGCACTTGTCTCAAGCTTTTAA 62

Qy 363 GlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGlnGlnAla 382
    |||||
Db 63 AACAGAAACCTCATCTCAATGAAGATGCTGATTAACCTTCCAGCAGCTTCAAGACT 122

Qy 383 LysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIle 402
    |||||
Db 123 AAGGTAAACACTTTAGAAACAGTGTCTCTTATGACTCTCTGCGAACCCTGAGATA 182

Qy 403 AspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGluCyArg 422
    |||||
Db 183 GACTTGTCTTGAAGAGGGCTCTCTGTTCACTTCTGTGGGAGACTTATAGTGCCGT 242

Qy 423 MetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPhe 442
    |||||
Db 243 TCGTGGTTGGCTTACAGTGAAGTACTCCCTTATAGAAATCATCTATGTAGAAATT 302

Qy 443 ValLeuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLysLeuLeu 462
    |||||
Db 303 GTCTTGAAAACTCAAGATGACATGACAGATCTTCCCGGCTTTGCAAACTGCTG 362

Qy 463 GluThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLysLysPhe 482
    |||||
Db 363 GAGGCATGGCTATATGAGAGTGTCTTCTAGATTAGAGACCCAAAGATATACATTC 422

Qy 483 LysLeuGlyAspTyrTyrAspAspPrometValLeuSerTyrLeuGluArgValGluVal 502
    |||||
Db 423 AACCTTGAGATTAATTATGATGATCTTACTGTCTGAGATATTATTAAGAGGCTTGAGGGA 482

```

QY 503 ValGInGlySerProLeuAlaAlaAlaTherMeAlaArgIleGlyAlaGlu----- 520
 |||||
 Db 483 GCAGGTGGTTCACCGCTTGGCTGACCGCGCTATATAGTGGATAGGTGAGAGGCTACT 542
 QY 521 -----HisVallyValAserAlaMetGlnAlaLeuGlnlyValPheProSer 536
 |||||
 Db 543 GCAGTTCCTTCATCATGTAAAGGCTAGTGAATTCAAGCATTTGCGAAGAGTTCCTCTT 602
 QY 537 ArgTyrThrAspArgAsnSerAlaGluProlyAspValGlnGluThrValPheSerVal 556
 |||||
 Db 603 CGTCGCTGAGAGAGAGAGCTGCCAGACATCAA----- 632
 QY 557 AspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAla 576
 |||||
 Db 633 -----TTAGATGTGTAGATGACAACTTCTT----- 659
 QY 577 ValArgProSerGluAsnPhenGluThr-----AsnAspTyrAlaAla 590
 |||||
 Db 660 -----CCTGTTGAAGTGAAGAGACTCTTGAAGAACTGATCAAGAGATTCGGCATC 713
 QY 591 ArgAlaGlyVal---SerGluSerSerValAspGluThrThrValGluMetSerValAla 609
 |||||
 Db 714 CTAGCTGAGGCTTCCTGGATATAGCAGTTTGAAGAGATGTATGAGAGAGAAACATATAGT 773
 QY 610 AspMetLeuIlyAspGluAlaSerValIlyIleLeuAlaAlaGlyValAlaIleGlyLeu 628
 |||||
 Db 774 GACAAATTTAAAGATGCCAGTGTCAAGATCATGTCTGTGTGTGTAATTGGAGTG 830

RESULT 11
 AI998415 561 bp mRNA linear EST 08-SEP-1999
 LOCUS AI998415.606 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
 DEFINITION thaliana cDNA clone 701545606, mRNA sequence.
 ACCESSION AI998415
 VERSION AI998415.1 GI:5845320
 KEYWORDS EST.

SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
 AUTHORS Arabidopsis thaliana (thale cress);
 Arabidopsis thaliana
 Burkhardt; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsys.
 1 (bases 1 to 561)

Chen, J., Montiyama, M., Chan, E., Mooney, M., Carrion, B., Gilliland, D.,
 Wang, X., Hillman, U., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
 Gorgone, G., Burns, D., Griffin, J., Mounoutoua, M., Nguyen, D.,
 Tan, R., Rose, M., Warren, B., Ton, B., Kaestury, K., Borillo, C.,
 Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
 Nobriza, A., Murry, L., Turner, C., Krikorian, S., Eider, L. and
 Hanson, D.

Arabidopsis thaliana Gene Expression Microarray
 Unpublished (1999)
 CONTACT: David Smoller, Ph.D.

Genome Systems, Inc., a wholly owned subsidiary of Incyte
 Pharmaceuticals, Inc.

4633 World Parkway Circle, St. Louis, MO 63134, USA

Tel: 877-577-2733

Fax: 314-427-3324

Email: service@genomesystems.com.

Location/Qualifiers
 1. 561

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone="701545606"

/tissue_type="rosette"

/dev_stage="4 - 7 weeks"

/clone_id="A. thaliana, Columbia Col-0, rosette-2"

/note="Vector: pSPORT, Site_1: NotI, Site_2: SalI, cDNA

library was derived from untreated rosette tissue from

Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.

Plants were grown in 1:1:1 peat moss/vermiculite/perlite

ORIGIN

Alignment Scores:

Pred. No.:	1,666-75	Length:	561
Score:	784.00	Matches:	157
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	3
Query Match:	19.3%	Indels:	0
DB:	1	Gaps:	0

US-10-600-070b-2 (1-801) x AI998415 (1-561)

QY 642 SerSerPheGlnArgIlyAspMetValSerSerMetGluSerAspValAlaThrIleGly 661
 |||||
 Db 559 TCATCTTTCAACGCAAGNATATAGTTTCTTCTATAGAACTGATGCTACCATAGGG 500
 QY 662 SerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrIleGlu 681
 |||||
 Db 499 TCAGTCAGAGCTGACGATTCAGAGCATTCCAGATGATGATGAGCTGACAGAGAA 440
 QY 682 IleValSerIlyerTrpGlnIlyIleIlySerLeuAlaPheGlyProAspHisArgIleGlu 701
 |||||
 Db 439 ATAGTATCCAGTGGCAGAGAGATTAAGTCTTGCTTGGGCTGATCACCGCATAGAA 380
 QY 702 MetLeuProGluValLeuAspGlyArgMetLeuIlyIleTrpThrAspAlaAlaGlu 721
 |||||
 Db 379 ATGTTACAGAGGTTTGGATGGCGAATGCTGAATTTGACGACAGAGCTGAA 320
 QY 722 ThrAlaGlnLeuGlyLeuValIlyAspTyrThrLeuLeuIlyLeuSerValAspSerVal 741
 |||||
 Db 319 ACTGCCAGCTGGGTGGTTGATATATGATTTACCTGTAAACTATCTGTGACAGTG 260
 QY 742 ThrValSerAlaAspGlyThrArgAlaLeuValAlaThrLeuGlnGluSerAlaCys 761
 |||||
 Db 259 ACAGTCTCAGCAATGAAACCCGCTCTGGTGGAGAGCACTGAGAGAGTCTGCTTGT 200
 QY 762 LeuSerAspLeuValHisProGluAsnAlaThrAspValArgThrTyrThrArg 781
 |||||
 Db 199 CTATCGATTTGGTTTATCATCAAGAAACATGACTGACATGCACTACACACAGAA 140
 QY 782 TyrGlnValPheTrpSerIlySerGlyTyrIlyIleThrGlnGlySerValLeuAlaSer 801
 |||||
 Db 139 TACGAAGTTTTCNGTCCAAAGTCAGGCTGGAATAATCATGAAAGCTCTGTTTGATCA 80

RESULT 12

BO834167

LOCUS

BO834167 534 bp mRNA linear EST 14-MAR-2003

DEFINITION A1EST0128 Arabidopsis lyrata inflorescence pMW-PCR library

Arabidopsis lyrata cDNA clone pMW1-D03 3', mRNA sequence.

BO834167

VERSION

BO834167.1 GI:28951482

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PubMed

COMMENT

Contact: Barrier M
 Department of Genetics
 North Carolina State University
 3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA
 Tel: 919 515 1761

Fax: 919 515 1695
Email: mbarrie@univ.ncsu.edu
Plate: 1 row: H column: 5
Seq primer: T3.

FEATURES

Source

Location/Qualifiers
1..534
/organism="Arabidopsis lyrata"
/mol_type="mRNA"
/cultiivar="Karhumaki"
/db_xref="taxon:59689"
/clone="PIW1-D03"
/tissue_type="Inflorescence"
/clone_lib="Arabidopsis lyrata Inflorescence pCMV-PCR library"
/note="Vector: pCMV-PCR (Stratagene); Created using PCR Library Construction kit (Stratagene)"

ORIGIN

Alignment Scores:

Pred. No.:	1,17e-74	Length:	534
Score:	776.00	Matches:	155
Percent Similarity:	95.7%	Conservative:	2
Best Local Similarity:	94.5%	Mismatches:	7
Query Match:	19.1%	Indels:	0
DB:	5	Gaps:	0

US-10-600-070b-2 (1-801) x BQ834167 (1-534)

QY 638 LeuLYSeSerSeSerSeSerPheGlnArgLYaSPMeValSeSerMeGluSeAspVal 657
DB 13 CTCAAAGACGACCTCATCTTTCAACGACGATATGAGCTTCTATGATCATGACATT 72
QY 658 AlaThrIleGluSerValArgAlaAspAspSerGluAlaLeuProArgMetAlaArg 677
DB 73 GCTACCATAGGCTCAGAGGACCTGATGATTCAGAACCTTCCCAATGATGCTAGG 132
QY 678 ThrAlaGluAsnIleValSerLYSTpGlnLYleYSerLeuAlaPheGluProAsp 697
DB 133 ACTGCAGAGGATAGTATGATCCAAATGCGACAAATCATCAAGCTTTGGGATGAT 192
QY 698 HisArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuLYleTPThrAsp 717
DB 193 CACTGATAGAAATGCTACAGAGGTTTGGATGGGAGATGCTGAAGATTGGAGTAC 252
QY 718 ArgAlaAlaGluThrArgIleGluValLYleValTYrAspTYrThrLeuLeuLYleuSer 737
DB 253 AGAGCTGTGGAACTGGCGAGCTTGGCTGTGTTATGATTATACACTGTGAACATATCT 312
QY 738 ValAspSerValThrValSerAlaAspGlyTYrArgAlaLeuValGluAlaThrLeuGlu 757
DB 313 GTTGAACGTGTGACAGTCTCAGACGATGAAACCGTCTGCTGAGAGCACTTGGAG 372
QY 758 GluSerAlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThr 777
DB 373 GAGTCTGCTGTATCTGATTTGGTTATCCAGAAACATGCTACAGATGTCAGAAC 432
QY 778 TYrThrThrArgTYrGluValPheTYrPserLYSerGlyTYrPlyLYleThrGluLYSer 797
DB 433 TACACACACAGATACGAAGCTGTGTGCTCAAGTCAAGGTCGAAATCACTGAAGGCTCT 492
QY 798 ValLeuAlaSer 801
DB 493 GTTCTGGATCA 504

RESULT 13
CN918842
LOCUS CN918842 703 bp mRNA linear EST 07-JUN-2004
DEFINITION 030207ABPB006290HT (ABPB) M9 root tips Malus x domestica cDNA clone
ABPB006290, mRNA sequence.
ACCESSION CN918842
VERSION CN918842.1 GI:48391655
KEYWORDS EST.
SOURCE Malus x domestica

ORGANISM

Malus x domestica

Epimorpha: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 703)
Bennings, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,
McArdney, S., Newcomb, R., Rose, G., Snowden, K., Walton, E. and Yauk, Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave, A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.

FEATURES

Source

Location/Qualifiers
1..703
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABPB006290"
/tissue_type="Root tips (distal 1.5 cm)"
/clone_lib=" (ABPB) M9 root tips"
/note="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

ORIGIN

Alignment Scores:

Pred. No.:	2,67e-74	Length:	703
Score:	774.50	Matches:	156
Percent Similarity:	74.9%	Conservative:	26
Best Local Similarity:	64.2%	Mismatches:	34
Query Match:	19.1%	Indels:	27
DB:	7	Gaps:	4

US-10-600-070b-2 (1-801) x CN918842 (1-703)

QY 1 MetGluAlaLeuSerHisValGlyIleGly-----LeuSerProPheGln 15
DB 14 ATGAAACCTTCMAACATTCGGCATCGCTTCTACATCCCGCTCTCCCTTCGCG 73
QY 16 LeuCYArg-----LeuProProAlaThrTYrLeuArgArgSerHisAsn 31
DB 74 CACCAACGAAACCCAAAGCTCTCTCC-----SerPheAlaThrAlaThrThr 103
QY 32 ThrSerThrThrIleCysSerAlaSerLYSTpAlaAspArgLeuSerAspPheAsn 51
DB 104 -----ACCATGCTTCGCGCAGCAATGCGGAGCGCTCTCGCGCATTCGCA 154
QY 52 PheThrSerAspSerSerSer-----SerPheAlaThrAlaThrThr 66
DB 155 TTCCTCGGCGACCT 214
QY 67 ThrAlaThrLeuValSerLeuProProSerIleAspArgProGluArgHisValProIle 86
DB 215 ACTCTGCT 274
QY 87 ProIleAspPheTYrGluValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIleArg 106
DB 275 CCCATCGATTTCTACACAGGTCGCGCCGCGCAGACATTTCTCTCGGAGCGGTATAGG 334
QY 107 ArgAlaPheGluAlaArgValSerLYSTpProProGlnPheGlyPheSerAspAlaLeu 126
DB 335 AGGCGCTGACGAGCTGAGGCTTCACAGCGCTTCAGTACGCTTACCCAGAGGCGCTG 394
QY 127 IleSerArgArgGlnIleLeuGlnAlaIleValCysGlyThrLeuSerAspProArgSerArg 146
DB 395 TTACGCCGAGGAGGATCTTCTAGCGCTTGGGAAACCTTACGCAACCCCGCTCACA 454
QY 147 ArgGluTYrAsnGluLYleuLeuAspArgGluAlaThrValIleThrAspValPro 166
DB 455 AGAGATGACACACAGAGCTTGGCGAAGAGGATGAAACCATCATCACTCAAGTTCTT 514

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2006, 12:44:39 / Search time 1193 Seconds
(without alignments)
4474.785 Million cell updates/sec

Title: US-10-600-070B-2

Perfect score: 4063

Sequence: 1 MEALSHVIGIGLSPFQCRLLP.....YEVFWKSGWKITEGVTLAS 801

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=X1h
-Q=/abs/ABSWEB.spool/US1060070/runat_21022006_112357_8602/app_query.fasta.1
-DB=N Geneseg -QMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAHSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abbs04
-USER=US1060070.QCGN_1_1727@runat_21022006_112357_8602 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21: *
1: geneseg1980s: *
2: geneseg1990s: *
3: geneseg2000s: *
4: geneseg2001as: *
5: geneseg2001bs: *
6: geneseg2002as: *
7: geneseg2002bs: *
8: geneseg2003as: *
9: geneseg2003bs: *
10: geneseg2003cs: *
11: geneseg2003ds: *
12: geneseg2004as: *
13: geneseg2004bs: *
14: geneseg2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4063	100.0	2406	12	ADJ38129	Adj38129 Arabidops
2	4051	99.7	2406	12	ADJ38208	Adj38208 Arabidops
3	4051	99.7	2637	12	ADJ38210	Adj38210 Plantid d
4	4051	99.7	2679	13	ADT14901	Adt14901 Plant cdm

ALIGNMENTS

RESULT 1	ID	ADJ38129	standard; cDNA; 2406 BP.
XX	XX	ADJ38129;	
AC	ADJ38129;		
DT	06-MAY-2004	(first entry)	
XX	XX		
DE	Arabidopsis thaliana Arc6-1 cDNA Seg1d.		
XX	XX		
KW	prokaryotic type; plastid division; FtsZ; ARC6; ARCS; Fzo; plant cell;		
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;		
XX	herbicide target; gene; ss.		
XX	XX		
OS	Arabidopsis thaliana.		
XX	XX		
PN	W02004001003-A2.		
PD	31-DEC-2003.		
XX	XX		
PF	20-JUN-2003; 2003WO-US019536.		
XX	XX		
PR	20-JUN-2002; 2002US-0390140P.		
PR	09-AUG-2002; 2002US-0402242P.		
XX	XX		
PA	20-JUN-2003; 2003US-00600070.		
XX	XX		
XX	(UNMS) UNIV MICHIGAN STATE.		

PI Oosteryoung KM, Vitcha S, Koksharova OA, Gao H;
 XX WPI: 2004-082486/08.
 DR P-PSDB; ADJ38202.

XX New isolated Ftn2, ARCS and/or Pzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.

XX Claim 1, SEQ ID NO 1, 287bp; English.

XX This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARCS), ARCS and Pzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.

XX Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2406
Score:	4063.00	Matches:	801
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-600-070b-2 (1-801) x ADJ38129 (1-2406)

QY 1 MetGluAlaLeuSerHisValIGlyLeuSerProPheGlnLeuCybAgluPro 20
 DB 1 ATGGAAGCTCTGAGTCACTGCGCATGCTCTCTCCCATTCATTCGCGATTACCA 60
 QY 21 ProAlaThrThrIleuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
 DB 61 CCGCGCAGCAAAAGCTCCAGCGTACGACACACCTCTACAACTATCTGCGCAGC 120
 QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnThrSerHisSerAspSerSerSer 60
 DB 121 AAATGGCCGACCTCTCTCTCGACTTCATTTCACTCCGATTCCTCTCTCTCC 180
 QY 61 PheAlaThrAlaThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 181 TTCGCCACCCGCCACACACCGCCTCTCTCTGCGCACCATCTATTGATCGTCCC 240
 QY 81 GluArgHisValProIleProIleAspPheTrgIleValLeuGlyAlaGlnThrHisPhe 100
 DB 241 GAAGCCAGCTCCCATCCCATTCATTCACGAGTATTAGAGCTCAAAACATTTTC 300
 QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProProGlnPheGly 120
 DB 301 TTAAACGATGGAATCAAAAGACATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGT 360
 QY 121 PheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
 DB 361 TTCAGCGACGACGCTTAATCAAGCGGACAGATTCCTCAAGCTGCTTCGAAACTCTG 420
 QY 141 SerAsnProArgSerArgArgGluTrpArgGlnGlyLeuLeuAspArgGluAlaThr 160
 DB 421 TCTAATCTCGGCTTAAGAAGAGATCAATGAAGTCTTCTGTGATGAAGAAAGCTACA 480
 QY 161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuCysValIleGlnGlnGly 180
 DB 481 GTCTATCTCATGTCTCTTGGAATAGAGCTTCGGGCTCTCTGTGATTCGAAAGAGT 540
 QY 181 GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuGlyGluArgProLys 200
 DB 541 GGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGGCTCTGCTTAAGAGAGGTTGCTTAAG 600

QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 601 TCGTTTAACCAAGATGTGGTTTAGTTATGGCGCTTCTCGATGTCTCGAAGGAT 660
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTrpGluPheValGluGlnAlaLeu 240
 DB 661 GCATATGGCAATGGATCCACTGATTTATTAAGTGTGTTAAGAGAAAGCTTTGG 720
 QY 241 LysLeuLeuGlnGluGluGluValAspSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
 DB 721 AAGCTTTTAACGAG 780
 QY 261 GluThrLeuGlnGluIleThrProArgTrpValLeuGlnLeuLeuGlyLeuProLeuGly 280
 DB 781 GAGACTTGGAAAGAGATCACTCGGCTTATGCTTGGACCTACTTGGCTTACCGCTTGGT 840
 QY 281 AspAspTrpAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 DB 841 GATGATTACGCTGCGAAAGACTAAATGGTTTAAAGCGGTGCGGAATATTTGTGTCT 900
 QY 301 ValGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
 DB 901 GTTGAAGAGGTGAT 960
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
 DB 961 GAGCGCTTTTTCAGATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAG 1020
 QY 341 AsnIleProAlaGluSerPheGluValTrpGluValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1021 AATATTCAGCAGAGATCATTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAG 1080
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGln 380
 DB 1081 TTATTTGTAAGAGACACCTTTTACAGATCTGATAGAGATTCAGCAACTTCAG 1140
 QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTrpAspThrArgAsnAsnTrp 400
 DB 1141 CAGGCTAAGGTATGCTATGAGATTCCTCGGAGTGTGTATGATACAGCAATTAATG 1200
 QY 401 GluIleAspPheGlyLeuGluLysGlyValLeuLeuIleGlyLysValAspGlu 420
 DB 1201 GAGATAGACTTCGGCTTAAGAAAGGAGCACTGTGCACTCTTAATAGCAAAAGTTGATGA 1260
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTrpArgAsnProAlaIleVal 440
 DB 1261 TGCCTGATGTGTGGGCTTACAGAGAGAGATTCACAAATATAGAAATCCAGCTATTGTG 1320
 QY 441 GluPheValLeuGlnAsnSerAsnArgAspAspAspAspLeuProGlyLeuCysLys 460
 DB 1321 GAGTTGTTTGGAGAAATTCAAATCGTATGACCAATGATCTCCCTGAGACTATGCAAA 1380
 QY 461 LeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
 DB 1381 TTGTTGAAAACCTGGTGGCAGGGGTTGCTTCTTCATAGGTTCAGAAACCAAAAGTTAA 1440
 QY 481 LysPheLysLeuGlyLysAspTrpTrpAspAspProMetValLeuSerTrpLeuGluArgVal 500
 DB 1441 AAATTTAACTCGGAGACTATGATGATGATTCCTATGTTTGAATCTTGGAAAGAGTG 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
 DB 1501 GAGGTAGTTCAAGGTTCTCTTATGCTGTGCTGCTGCAACTATGCAAGATGAGCGAG 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTrpThrAsp 540
 DB 1561 CATGTAAAGCTAGTGTCTATGACGACCTGCAAGAAAGTTTCTCCCTTACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
 DB 1621 AGAACTCGGCTGAAACCCCAAGAGATGCAAGAGACAGTGTATTAGTATGATCTGTGTGT 1680

QY 561 AsnAnValGlyATGAPGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1681 AACATGATAGGCGCGTATGCGAGCCCTGGTCTTTATTCAGAAAGCTGTAAGCCCTCT 1740
 QY 581 GluAsnPheGluThrAsnAPTYrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1741 GAAAACTTGGAAATGATATGATATGCAATTCAGCTGGGGCTCAGAGAGTACGGTTGAT 1800
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuysGluAlaSerValIleLeu 620
 DB 1801 GAAACTGCTGTTGAAAGTCCGGTGTGATATGTTAAAGGAGGCAAGGTGAACACTTA 1860
 QY 621 AlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGluNlyTYrPheLeuIleSer 640
 DB 1861 GCTGCTGCTGGGCAATTCGATTCATTCCTGTCAGCCGAAAGTATTTCTTAAAGC 1920
 QY 641 SerSerSerPheGluArgIleAspMetValSerSerMetGluSerAspValAlaThrIle 660
 DB 1921 AGCTCATCTTTTCAACCAAGATATGGTCTTCTATGGAATCTGATGCTACCATTA 1980
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 1981 GGGTCAGTCAAGCTGACGATTCAGAAACATTCCTCCGAATGATGCTAGAGCTGCAGAG 2040
 QY 681 AsnIleValSerIleTYrGluNlyIleIleSerLeuAlaPheGlyProAspHisArgIle 700
 DB 2041 AATATGATATCCAAAGTGGCAGAAATTAAGCTCTGGCTTTGGGCTGATCACCAGATA 2100
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuysIleTYrThrAspArgAlaIle 720
 DB 2101 GAAATGTTACCAAGGTTTGGATGGGCGAATGCTGAATTTGGAAGTGCAGAGCAGCT 2160
 QY 721 GluThrAlaGluLeuGlyLeuValTYrAspTYrThrLeuLeuysLeuSerValAspSer 740
 DB 2161 GAAACTGCGAGCTGGTGGTGGTATTAATTAACATCTGTTGAAACATCTGTTGACAGT 2220
 QY 741 ValThrValSerAlaAspGlyTYrThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 760
 DB 2221 GTGACATCTTCAGAGATGAAACCCGTGCTGTGTGAAAGCACTCTGGAGAGCTGTGCT 2280
 QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTYrThrThr 780
 DB 2281 TGCTATCTGATTTGGTTCAATCCAGAAACAAATGCTACTGATGTCAGAACCTACACACA 2340
 QY 781 ArgTYrGluValPheTYrPheSerIleSerGlyTYrIleValIleThrGluGlySerValLeuAla 800
 DB 2341 AGATACGAAGTTTCTGTCTCAAGTCAAGGTGGAATACTGAAGGCTCTGTCTTGCA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403
 RESULT 2
 ID ADJ38208 standard; DNA; 2406 BP.
 AC ADJ38208;
 DT 06-MAY-2004 (first entry)
 XX
 DE Placid division-related Arc6 orthologue gene 2.
 XX
 KW prokaryotic type; placid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 XX herbicide target; gene; de.
 OS Arabidopsis thaliana.
 PN WO2004001003-A2.
 PD 31-DEC-2003.
 PF 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 PI Oseeryoung KW, Vicha S, Koksharova OA, Gao H;
 XX WPI, 2004-082486/08.
 DR P-PSDB; ADJ38207.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing placid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.
 PS
 CC This invention relates to novel prokaryotic type or placid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterising placid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.
 SQ
 SQ Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 2406
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-600-070B-2 (1-801) x ADJ38208 (1-2406)
 QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGluLeuCyArgLeuPro 20
 DB 1 ATGGAAGCTTGATGACGATCGGCAATTGCTCTCCCATTCATTCATTCATTCATTCATTC 60
 QY 21 ProAlaThrThrIleValLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
 DB 61 CCGGCGACGACAAAGCTCCGACGTAAGCCACAAACCTCTCAACTATCTCTCCGACG 120
 QY 41 LysThrAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 DB 121 AAATGGGCGCAGCGTCTCTCTCGAATTAAATTCACCTCCGATTCCTCTCTCTCTCC 180
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 181 TTGCGCACCGCCACACACCGCAGCTCTGCTCTCCGACACATTAATGATCGGCC 240
 QY 81 GluArgHisValProIleProIleAspPheTYrGluValIleGlyAlaGluIleHisPhe 100
 DB 241 GAAAGCCACGTCCTCCCATCCCATTTGATTTCTACAGGATTAAGAGCTCAAAACATTC 300
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleProProGluIlePheGly 120
 DB 301 TTAAACCATGAAATTCAGAAAGCAATTCGAAAGGTTCGAAACCGCCGCAATTCGGT 360
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGluIleGluGluAlaIleCysGluThrLeu 140
 DB 361 TTCAGCGACGACCGCTTAAATCAACCGGAGACAAATTCCTCAACTGTTCCGAAACTCTG 420
 QY 141 SerAsnProArgSerArgArgGluTYrArgGluGlyLeuLeuAspAspGluGluAlaThr 160
 DB 421 TCTAATCTCTGCTAGAGAGATCAATGAAGGCTCTTCTGATGATGAAGAGCTACAA 480

QY 161 ValIleThraPseValProTTPAAsPLeuValProG1ValLeuCyValLeuGlnGluGly 180
 DB 481 GTATCACTGATGTTCTTGGATTAAGTTCTCTGCTCTCTCTGATTAAGCAAGAGT 540
 QY 181 GlyGluThrGluIleValLeuAArgValGlyGluAlaLeuLeuGlyGluArgProGly 200
 DB 541 GGTAGACCTGAGATGATGTTCTTGGGTTGTGAGGCTCTGCTTAAGGAGAGTTCCTAG 600
 QY 201 SerPheLeuGlnPseValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 601 TCGTTTAAGCAAGATGTTGTTTATGATGAGCGCTTCTGCTTCTGATTCGAGGAGAT 660
 QY 221 AlMetAlaLeuAspProPsePheIleThrGlyTyrgIuPheValGluGluAlaLeu 240
 DB 661 GCTATGCACTTGCATCCACTGATTTTAACTGTTATGAGTTTGTGAGGAAGCTTGTG 720
 QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuAArgAlaGlnIleAsp 260
 DB 721 AACCTTTTACAGAGAGAGAGAGCAAGTACCTTGCACCGGATTTACCTGCACAATTTGAT 780
 QY 261 GluThrLeuGlnGluIleThrProArgTyrgValLeuGlnLeuGlnGlyLeuProLeuGly 280
 DB 781 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGTGAGCTACTTGCCTTACCGCTTGT 840
 QY 281 AspAspTyrgAlaAlaValArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 DB 841 GATGATTAACCTCGGAAAGACCTAAATGTTTAAAGCGTGTGCGGAATATTTGTGCTCT 900
 QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThraArgGluLysPheMetAsn 320
 DB 901 GTTGAGAGAGGTGAGATCAGCTCTTGTGGGGTGTGACCCCTGAGAAATTTATGAT 960
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAlaPheLeuPheValAlaThraProSer 340
 DB 961 GAGCGCTTTTACAAATGACAGCTGCTGAGAGAGGTATCTTTTGTAGCTAACCCCAAGC 1020
 QY 341 AsnIleProAlaGlnSerPheGluValTyrgIuValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1021 AATATTCACAGAGATCTTGAAGTTTAAAGTTTCAAGTTGCACTGCTTGTGGCTCAAGCT 1080
 QY 361 PheIleGlyLysIysProHileLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1081 TTTATTTGTAAGAGACACACCTTTTACAGATGCTGATTAAGCAATTCACAGCACTTCAG 1140
 QY 381 GluAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrgAspTrpArgAspAspTrp 400
 DB 1141 CAGCTTAAGGTAAATGCTATGAGATTCCTGCAATGTATGATACAGGAATTAATTTGG 1200
 QY 401 GluIleAspPheGlyLeuGlnArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
 DB 1201 GAGATTAAGCTTCGCTCTAGAAAGGAGCTCTGTCACTCTTATAGCCAAAGTTGATGAA 1260
 QY 421 CysArgMetTrpLeuGlnLeuAspSerGluAspSerGlnTyrgAspProAlaIleVal 440
 DB 1261 TGCCTATGCTGCTGCTTACAGTACAGTATCAAAATTAAGAAATTCACGCTATTTGTG 1320
 QY 441 GluPheValLeuGlnuAsnSerAsnArgAspAspAspAspLeuProGlyLysCysLys 460
 DB 1321 GAGTTTGTGTAAGAAATTCAAATCGTATGATACATATGATTCCTCGAGCTATGCAAA 1380
 QY 461 LeuLeuGlnThraTrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
 DB 1381 TTGTTGAAACCTGTTGGCGAGGCTTCTTCTTCTAGTTTCAAGACACCAAGATTA 1440
 QY 481 LysPheLysLeuGlnAspTyrgTyrgAspAspProMetValLeuSerTyrgLeuGlnArgVal 500
 DB 1441 AAATTTAACTCGGAGACTACTATGATGATCTCTATGCTTTGAGTTTACTTGAAGAGTGT 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaAlaThraMetAlaArgIleGlyValGlu 520
 DB 1501 GAGGTAGTTTCAAGGCTTCTCTTCTTCTGCTGCTCAAGTATGAGCAAGATTTGAGCCGAG 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrgThraAsp 540

DB 1561 CATGTAAAGCTAGTACTATGACAGGCACTGCACAAAGTTTTCCTCCCTATACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThraValPheSerValAspProValGly 560
 DB 1621 AGAACTCGGCTGAAACCAAGAGTGTGAGAAACAGCTGTTTATGATGATCTGTTGCT 1680
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValAlaPheIleAlaGlnAlaValArgProSer 580
 DB 1681 AACATGTAGGCGGTATGATGATGAGCTCGGTCTTTATTCAGAACTGTAAAGCTCTCT 1740
 QY 581 GluAsnPheGlnThraAspTyrgAlaIleArgAlaGlyValSerGlnSerSerValAsp 600
 DB 1741 GAAACTTTGAAACTATGATTTATGCAATTCGAGTGGGCTCTCAAGAGTACGTTGAT 1800
 QY 601 GluThraValGlnMetSerValAlaAspMetLeuLeuGlnAlaSerValLysIleLeu 620
 DB 1801 GAAACTACCTGTTAAATATGCTCGTTCGATATTTAAAGGAGCAAGTGTAAAGATCTTA 1860
 QY 621 AlAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrgPheLeuLysSer 640
 DB 1861 GCTGCTGTGTGCAATTTGACATTTTCACTGTTCAAGCAGAAATATTTCTTAAAGC 1920
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAspValAlaThrIle 660
 DB 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgTrpAlaGlu 680
 DB 1981 GGGTCAGTCAAGCTGACATTCATCAAGAGCACTTCCAGATGAGATGCTAGACTGCAGAG 2040
 QY 681 AsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHileArgIle 700
 DB 2041 AATATGATTCCAAGTGTGAGAGATTAAGTCTCTGCTTTTGGCTGATTCACCGATA 2100
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTrpThraAspAlaAla 720
 DB 2101 GAAATGTTACCAAGGTTTGTGAGTGGCGAATCTGAGATTTTGGACTGACAGAGCAGCT 2160
 QY 721 GluThraAlaGlnLeuGlyLeuValTyrgAspTyrgThraLeuLysLysSerValAspSer 740
 DB 2161 GAACCTCGGAGCTGTGGTGTGTTTATGATTAACATGTTGAAACTATCTGTTGACAGT 2220
 QY 741 ValThraValSerAlaAspGlyThraArgAlaLeuValGluAlaThraLeuGlnGlnSerAla 760
 DB 2221 GTGACACTCTCACAGATGAGAACCCGCTCTGTGTGAAACCACTCGAGAGTCTGCT 2280
 QY 761 CysLeuSerAspLeuValHisProGlnuAsnAlaThraAspValArgThrTyrgThraThr 780
 DB 2281 TGTCTATCTGATTTGTTTCACTCCAGAAACAAATGCTATGATGTCAAGAACCTTACACA 2340
 QY 781 ArgTyrgIuValPheTrpSerLysSerGlyTrpLysIleThraGlnGlnGlnSerValLeuAla 800
 DB 2341 AGATACCAAGTTTGTGTGCTCAAGTCAAGGTGAAATATCTAAGAGCTCTGTTCTTGCA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403

RESULT 3
 ADJ38210
 ID ADJ38210 standard; DNA; 2637 BP.

AC ADJ38210;
 DT 06-MAY-2004 (first entry)
 DE Plasmid division-related Arc6 orthologue gene 3.

XX prokaryotic type; plasmid division; Fnc2; ARG6; ARCS; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; gene; ds.

OS Arabidopsis thaliana.
 XX MO2004001003-A2.
 XX 31-DEC-2003.
 XX 20-JUN-2003; 2003MO-US019536.
 XX 20-JUN-2002; 2002JUS-0390140P.
 PR 09-AUG-2002; 2002JUS-0402242P.
 PR 20-JUN-2003; 2003JUS-00600070.
 XX (UNMS) UNIV MICHIGAN STATE.
 PA
 PI Osteoyoung KM, Vitsha S, Koksharova OA, Gao H;
 XX WPI; 2004-082486/08.
 DR P-PSDB; ADJ38209.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 PS
 PS Disclosure; Fig 8; 2877p; English.
 XX
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (AR6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.
 XX
 SQ Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 2637
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-600-070b-2 (1-801) x ADJ38210 (1-2637)
 QY 1 MetGluAlaLeuSerHisValGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 DB 114 ATGGAGCTCTGAGTCACGTCGGCATTTGCTCTCCCATTCCTCAATTATCCGATTACCA 173
 QY 21 ProAlaThrXlybLeuArgSerHisAsnThrSerThrXlleCySerAlaSer 40
 DB 174 CCGGCACGACAAAGCTCCGACGTCACCAACACCTCTACAACTATCTCTCCGACG 233
 QY 41 LysTTPAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSer 60
 DB 234 AATATGGCCGACCGCTCTCTCTCCGACTTCAATTTCACCTCCGATTCTCTCTCTCC 293
 QY 61 PheAlaThrAlaThrThrAlaThrAlaThrValSerLeuProProSerIleAspArgPro 80
 DB 294 TTGCGCACCGCACCAACGACGCTCTCTCTCCGACCACTTATTTGATCGTCC 353
 QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100
 DB 354 GAACGCAAGCTCCCATTCCTCCATTTTCTACAGGATTATGAGAGCTCAACATTTCC 413
 QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProGlnPheGly 120
 DB 414 TTAAACGATGAGATTCAGAAAGACATTCGAAGCTTATGGGTTTCAAAACGCGCAATTCGGT 473
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCySGlnThrLeu 140

DB 474 TTCAGGACGACGCTTTAATCAAGCCGAGACAAATCTTCAAGCTGTTGGAACCTCTG 533
 QY 141 SerAsnProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspArgGlnAlaThr 160
 DB 534 TCTAATCTCTCGCTCTAGAGAGAGTACAAATGAAGCTCTTCTGATGATGAAGAGCTTACA 593
 QY 161 ValIleThrAspValProTTPAspLysValProGlyValLeuCybValLeuGlnGlnGly 180
 DB 594 GTCACTACGATGTTCTTGGAAAGATTCCTGGTCTCTGATGATGATGATGATGATGATGAT 653
 QY 181 GlyIleThrGlnIleValLeuArgValGlyGlyAlaLeuLeuIleGlyArgLeuProLys 200
 DB 654 GGTGAGACTGAGATGATGTTCTTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 713
 QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 714 TCGTTTAAAGCAAGATGCTTTAGTTATGCGCTTCTCGATGCTCGAGCGAT 773
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGlnPheValGlnGlnAlaLeu 240
 DB 774 GCTATGCACTGATGATCCACTGATTTTAACTGTTATGATGATGATGATGATGATGATGAT 833
 QY 241 LysLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
 DB 834 AACCTTTACAG 893
 QY 261 GlnThrLeuGlnGlnIleThrProArgTyrValLeuGlnLeuLeuGlyLeuProLeuGly 280
 DB 894 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTTACCTTGGCTTACCGCTTGGT 953
 QY 281 AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTTPSer 300
 DB 954 GATGATTTAGCGCTCGCAAAAGACATTAATGTTTAAAGCGTGTGGAGATATTTTGCTCT 1013
 QY 301 ValGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGlnLysPheMetAsn 320
 DB 1014 GTTGGAGAGAGTGGAGACATCACTCTTGTGGGGGTTTGAACCGTGAGAAATTTATGAAT 1073
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
 DB 1074 GAGGCTTTTTCAGATTCAGATTCAGCTGTCGAGAGGTTGATCTTTTTCAGCTTACCCCAAGC 1133
 QY 341 AsnIleProAlaGlnSerPheGlnValTyrGlnValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1134 AATATTCACAGACAGATTCATTTGAAGTTTACGAAGTTGCACTTGTGGCTCAAGCT 1193
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1194 TTTATTTGTTAAGAGCAACCTTTTACAGATGCTGATAGCAATTCACAGCAATTCACAG 1253
 QY 381 GlnAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAsnTTP 400
 DB 1254 CAGGCTTAAGATATGCTATGAGATTCCTGCGCATGTTGATGATGATGATGATGATGATGAT 1313
 QY 401 GlnIleAspPheGlyLeuGlnArgGlyLeuCybAlaLeuLeuIleGlyLysValAspGln 420
 DB 1314 GAGATGACCTTCGCTTAAAGAGGAGCTCTGCACTGCTTATAGCAAAAGTTGATGAA 1373
 QY 421 CysArgMetTTPLeuGlyLeuAspSerGlnAspSerGlnTyrArgAsnProAlaIleVal 440
 DB 1374 TGCCGATGTTGTTGGCTTAAAGATGAGATGATGATGATGATGATGATGATGATGATGAT 1433
 QY 441 GluPheValLeuGlnLysSerHisAsnArgAspAspAspAspLeuProGlyLeuCybLys 460
 DB 1434 GAGTTTGTGTTGGAGATTCAAATCGTATGACAAATGATGATGATGATGATGATGATGATGAT 1493
 QY 461 LeuLeuGlnThrTTPLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
 DB 1494 TTGTTGAAACCTGTTGGCAAGGCTTCTTCTTCAAGGTTCAAGAACCAAAAGTTAA 1553
 QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGlnArgVal 500

Alignment Scores:

Pred. No.: 0 Length: 2679
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 13 Gaps: 0

US-10-600-070b-2 (1-801) x ADT14901 (1-2679)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 DB 115 ATGGAACCTCGAAGTCAAGTCGGCATTTGCTCTCCCATTCACATTATATGCGATTACA 174
 QY 21 ProAlaThrThrIleValLeuArgArgSerHisAlaThrSerThrTrpIleCysSerAlaSer 40
 DB 175 CCGCGGACGACAAAGCTCCGACGTAGCCAAACACTCTACAACTATCTGCTCCGACAGC 234
 QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 DB 235 AATGGGCGGACGCTCTTCTCTCCGACTTCACATTTCACCTCCGATTCTCTCTCTCC 294
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 295 TTGGCCACCGGACACCAACCGGACCTCGTCTCTCGGACCATGATTGATCGTCCC 354
 QY 81 GluArgHisValProIleProIleAspPheTrpGlnValLeuGlyValGlnThrHisPhe 100
 DB 355 GAACGCGACGCTCCCATCCCATTTGATTTTCAACAGGATTTAGAGCGCTCAAAACATTTTC 414
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleValProProGlnPheGly 120
 DB 415 TTAAACGATGAGATCAAGAGAGCATTCAGAGCTAGGCTTGGAAACGCGCCCAATTCGT 474
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
 DB 475 TTGACGACGACGCTTTATATCAGCGGACAGATTCTTCAAGCTGCTGGGAAATCTGT 534
 QY 141 SerAsnProArgSerArgArgGluTrpAsnGluGlyLeuLeuAspAspGluAlaThr 160
 DB 535 TCTAATCTCGCTGTACAGAGAGCATCAATCAATGAGCTTCTTGATGATGAAGAGCTACA 594
 QY 161 ValIleThrAspValProTrpAspIleValProGlyAlaLeuCysValLeuGlnGluGly 180
 DB 595 GTATCATCTGATGTTCTCTGGAGTAAGGTTCTGTGCTCTCTGTGATTCAGAAAGGT 654
 QY 181 GlyIleThrGluIleValLeuArgValGlyValAlaLeuLeuGlyGluArgProIle 200
 DB 655 GGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGCTCTGCTTAAGAGAGGTTGCCCTAG 714
 QY 201 SerPheIleGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 715 TCGTTTAAAGCAAGATGTGTTTGTATGCGCTTCGTTCTCGATGCTCGAGGAGAT 774
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTrpGluPheValGluGluAlaLeu 240
 DB 775 GCTATGCAATTTGATTCACCTGATTTTAACTGATTAAGATTGTTGAGAAAGCTTTG 834
 QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
 DB 835 AAGCTTTTACAG 894
 QY 261 GluThrLeuGluGluIleThrProArgTrpValLeuGluLeuLeuGlyLeuProLeuGly 280
 DB 895 GAGCTTTGAGAGAGATCACTCCGCTTATGCTTGTGAGCTACTTGCTTACCTTGTGT 954
 QY 281 AspAspTrpValAlaIleValArgLeuLeuGlyLeuSerGlyValAlaGlyAlaIleLeuTrpSer 300
 DB 955 GATGATTAACCTGCGAAAGAGCTAAATGTTTAAAGCGGTGCGGAAATTTTGTGTCT 1014
 QY 301 ValGlyValGlyAlaSerAlaLeuValGlyValLeuThrArgGluValPheMetAsn 320
 DB 1015 GTTGGAGAGGTGAGAGATCACTCTTGTGGGGGTTTGAACCCGTGAAGAGTTTATGAT 1074

QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
 DB 1075 GAGCGGTTTTTACAAAGACAGCTGCGAGCGGTTGATCTTTTGTAGTACCCCAAGC 1134
 QY 341 AsnIleProAlaGluSerPheGluValTrpGluValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1135 AATATTCAGAGAGATCATTTGAAGTTTACGAAGTTCACCTGCTTGTGGCTCAAGCT 1194
 QY 361 PheIleGlyValIleProIleLeuLeuGlnAspAlaAspIleGlnPheGlnLeuGln 380
 DB 1195 TTTATTTGTAAGAGACCACTTTTACAGAGTCTGATTAAGCAATTCAGCACTTCAG 1254
 QY 381 GluAlaIleValMetAlaMetGluIleProAlaMetLeuTrpAspThrArgAspAsnTrp 400
 DB 1255 CAGGCTAAGGTAATGCGTATGAGATTCCTGCGAGTTGTATGATACCGGAAATTAATGG 1314
 QY 401 GluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyValAspGlu 420
 DB 1315 GAGATACATTCGCTTGAAGAGGAGACTGTGCTCATCTGCTTATAGGCAAGTTGATGAA 1374
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTrpArgAsnProAlaIleVal 440
 DB 1375 TGCCGTATGTGTGGCTTAGACAGTGAAGATTCAAAATATAGAAATCCAGCTATTGTG 1434
 QY 441 GluPheValLeuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLeu 460
 DB 1435 GAGTTGTTTTTGAAGAAATCAANTCGTAAGACATATATATCTCCCTGACATATGCAA 1494
 QY 461 LeuLeuGluIleThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrIleAspIle 480
 DB 1495 TTGTTGAAACCTGCTGTGGACGGGTTGCTTCTTCTAGCTTCAGAGACCAAAAGATTA 1554
 QY 481 LysPheIleLeuGlyAspTrpTrpAspAspProMetValLeuSerTrpLeuGluArgVal 500
 DB 1555 AATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGTAGTTACTTGAAAGAGTG 1614
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyValGlu 520
 DB 1615 GAGGTAGTCAAGGTTTCTCTTATGCTGCTGACCTATGGCAAGATTTGGAGCGGAG 1674
 QY 521 HisValIleValAlaSerAlaMetGlnAlaLeuGlnIleValPheProSerArgTrpAsp 540
 DB 1675 CATGTGAAGCTAGTGTCTATTCAGAGCACTGCAAGAAATTTTCTTCCGCTATACGAT 1734
 QY 541 ArgAsnSerAlaGluProIleAspValGlnGluThrValPheSerValAspProValGly 560
 DB 1735 AGAACTCGGCTGAACCAAGAGATGTCAAGAGACAGTGTATAGTATCCTGTTGT 1794
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1795 AACATATAGGCGGTGATGTGAGCCGTGGTCTTTATTCAGAAAGCTGTAAGCCCTCT 1854
 QY 581 GluAsnPheGluThrAsnAspTrpAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1855 GAATACTTTAAACTAATATGATTTATGCAATTCAGACTGGGCTTCAGAGAGTACGTTGAT 1914
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuGlyAlaSerValIleValLeu 620
 DB 1915 GAATCATCTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1974
 QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleTrpPheLeuIleSer 640
 DB 1975 GCTGCTGTGTGCAATTTGACATTCATCTGTCAGGAGAGATTTTCTTTAAAGC 2034
 QY 641 SerSerSerPheGlnArgIleAspMetValSerSerMetGluSerAspValAlaThrIle 660
 DB 2035 AGCTCATCTTTTCAACGCAAGAGATATGTTTCTTCTATGAAATCTGATGTCGCTACATA 2094
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 2095 GGGTCAGTCAAGAGCTGACATTCAGTTCAGAAAGCACTTCCGAGATGATGTGAGACTGAG 2154

Oy	661	AenliVeIseVleTRpGlnlysllelysserleuAlapheGlyProAspHiSaRgtIle	700
Db	2155	AATATAGATCCAAAGTGGCAGAAATTAAGTCTCTGGCTTTTGGGCTGATCACCGCAT	2214
Oy	701	GIuuectleuProGluValleuAbpGlyArqmetleuVsiletrpThrAspArAlaA	720
Db	2215	GAAATGTTACCAAGGTTTTGGATGGGGAATGCTGAAGATTGGACTGACAGAGCAgCT	2274
Oy	721	GIuThralaGlnleuGlyleuValTyraApTyThrleuLeuylbLeuSerValApbSer	740
Db	2275	GAAATCGCGCAGCTTGGGTGGTTATATATACACTGTGAACCTATCTGTTGACgCT	2334
Oy	741	ValThrValSerAlaApbGlyThrArpAlaleuValGluAlaThrleuGluGluSerAla	760
Db	2335	GTGAAGTCTCAGCAGATGAACCCGTCCTCTGGTGGAAAGCACTTGGAGAGTCTGCT	2394
Oy	761	CysLeuSerAspLeuValHIsProGluAbnAbnAlaThrAspValArgThrTyThrThr	780
Db	2395	TGTCATCTGATTTGGTTCATCCAAAACAATGTAATGTCAGAACTACACAA	2454
Oy	781	ArgTyrgluValPheTyrPserIysSerGlyTTrpIyVllethrgluGlySerValleuAla	800
Db	2455	AGATTCAGAAATTTCTGGTCCAAGTCAAGGTGGAATAATCACTGAAGGCTCTGTTCTTGA	2514
Oy	801	Ser 801	
Db	2515	TCA 2517	
RESULT 5			
ADJ38135			
ID	ADJ38135	standard; cDNA; 2406 BP.	
AC	ADJ38135;		
XX			
DT	06-MAY-2004	(first entry)	
XX			
DE	Arabidopsis thaliana AtFtn2 cDNA SeqID9.		
XX			
KM	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;		
KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;		
KM	herbicide target; gene; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	MO2004001003-A2.		
XX			
PD	31-DEC-2003.		
XX			
PF	20-JUN-2003; 2003MO-US019536.		
XX			
PR	20-JUN-2002; 2002US-0390140P.		
PR	09-AUG-2002; 2002US-0402242P.		
PR	20-JUN-2003; 2003US-00600070.		
XX			
PA	(UNMS) UNIV MICHIGAN STATE.		
XX			
PI	Osereyoung KM, Vitha S, Koksharova OA, Gao H;		
XX			
DR	MP1; 2004-082486/08.		
XX	P-P8DB; ADJ38203.		
XX			
PT	New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful		
PT	for further characterizing plastid division in plant cells, and in		
PT	varying agronomic and horticultural characteristics of economically		
PT	important plants.		
XX			
PS	Claim 1; SEQ ID NO 9; 287pp; English.		
XX			
CC	This invention relates to novel prokaryotic type or plastid division and		
CC	related genes and proteins. In particular, the invention relates to novel		
CC	Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and		
CC	compositions of the present invention are useful for further		
CC	characterising plastid division in plant cells, in order to vary		

agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is a cDNA sequence which is related to the invention.

Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2
Score:	4049.00	Matches:	7
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	2
Query Match:	99.7%	Indels:	0
DB:	12	Gaps:	0

US-10-600-0708-2 (1-801) X ADJ38135 (1-2406)

QY I MetGUALaleuSerHisValGIlyIleGIyleuSerProPheGlnIleuCysArgLeuPro 20
Db 1 ATGGAAGCTGTGAGTACGTCGGCATTTGGTCTCTCCCATTCATTATGCCAATTACA 60

QY 21 ProIaIaThrIyLysLeuArgAysSerHisAsnThrSerThrThrIleCysSerAlaSer 40
61 CCGGAGCAGCAAAAGCTCCAGCGTAGGCACAAACCTCTACAACTATCTGCTCGGCAGC 120
Db

QY 41 LysTPAlaspArgLeuleuSeraspPheahnpHeThiserAspserserserserser 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AATGGGCGCAGCTTCTCTCCGACTTCATTCACTTCGATTCTCTCTCTCTCC 180

```

Oy      61 PhealathralathrThrThralathrLeuValSerleuProPserIleaspArgPro 80
        |||||
Db      181 TTGGCACCAGCACCACGCGCACTCTGCTCTGTGCAACATCTATTGATCGTCCC 240

```

```
QY      81  GIuAghIvAlProIleProlleapPheTyGInValLeuGIyAlaGIInThrHisPhe 1000
      |||||
Db    241  GAAGCCACGTCCTCCCATTCGATTGTTTCAACAGATATTAGAGCTCAACACATTTTC 3000
```

QY 101 LeuThrAspGlyIleArgArgAlaPheGluIleArgValSerLysProProGlnPheGly 120
| | | | |
Db 301 TTAACCGATGAATCAGAAGCATTCGAAGCTTAGGGTTTGAAACGCCGCATTCGGT 360

Qy 121 pHeSerAspApAlaLeuIleSerArgGlnIleuGlnAlaAlaCysglutThrIeu 140
|||
Db 361 TTCAGCGACGACGGTTAATCAAGCCGAGACAGATTCTTCAAGCTGCTTGCAAACTGTG 420

QY 141 SerAsnProArgSerArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
|||
Db 421 TCTATCCTCGGTCTAGAGAAGTCAATGAAGCTTCCTTGATGATTGAAAGACTACA 480

QY 161 ValIeThraSpvaIProTPAspyVaAlProGIyAlaleuCysVaAlleugInglucly 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 GTCAACACTAATGTTCTTGSGATAAGTGTCTCGGGCTCTGTGTATTCAGAAGGT 540

QY 181 GlyGluThrGutIleValLeuArgValGIysAlaIeuleuLysGIuArgIeuProlys 200
|||
DB 541 GGTGACACTGATAGTCTTCGGGTTGGTAGGCCTCGCTTAAGAGAGAGTTGCCCTAAG 600

Qy 201 SerPheLySGlnAspValValLeuValMetAlaLeuAlaPheLeuaspValSerArgASP 220
| | | | |
Db 601 TCGTTTAAGCAGAAGATGCTTATTAGTGATGGCGCTTCGCATGTCTCGAGGGAT 660

Oy 221 ALawetAlaleuAspProPcoaspheilethrGlyTYrGlupheValGluGuaIaleu 240
Db 661 GCTATGGCATTGCATCCAGTATTATTAACGTGTTATGATGATTGTGAGAGAAGCTTTC 720

QY 241 LysLeuLengInGluGluGlyAlaSerSerIeuaLPProAbPleuAgaInIleap 26C
|||
Db 721 AAGCTTTACGAGAGAAGGCAACTAGCCTGCACCGAAATTACGTGCACAATTCAT 780C

[illegible]

QY 281 AspAspTyrAlaAlaValSerLeuAsnGlyLeuSerGlyValAlaGluAsnIleLeuTyrSer 300
 DB 841 GATGATTACCGCTCGGAAAGACCTAAATGGTTTAAAGCGGTGGCGAAATATTTTGTGCT 900
 QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyLeuThrArgGlyLysPheMetAsn 320
 DB 901 GTTGAAGAGGTGAGCATCAGCTCTTGTGGGGGTTGAACCGGTGAAGATTATGAAAT 960
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlyValAspLeuPheValAlaThrProSer 340
 DB 961 GAGCGCTTTTATGAATGACAGCGCTGACACAGGTTTCATCTTTTGTAGCTACCCCAAGC 1020
 QY 341 AsnIleProAlaGluSerPheGluValTyrGlyValAlaLeuAlaLeuValAlaGluAla 360
 DB 1021 AATATTCACAGAGTCAATTTGAAGTTTACAGAGTTGCACTTGTCTTGGCTCAAGCT 1080
 QY 361 PheIleGlyLysProAlaLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1081 TTTATTGGTAAGAACCCACACTTTTACAGGATGCTATTAAGCAATTCAGCAACTTCAG 1140
 QY 381 GlnAlaLysValMetCalaMetGluIleProAlaMetLeuTyrAspThrArgAsnMetTyr 400
 DB 1141 CAGGCTAAAGGTAAATGGCTATGAGATTCCTGCGATGTTGATGATACAGCAATAAATGG 1200
 QY 401 GluIleAspPheGlyLeuGluValArgGlyLeuGlyAlaLeuLeuIleGlyLysValAspGlu 420
 DB 1201 GAATATGACTTCGGTCTGAAAGAGGAGCTGTGCTGCTATTAAGCAAGTTGATGAA 1260
 QY 421 CysArgMetTyrLeuGlyLeuAspSerGlyLysAspSerGlnTyrArgAsnProAlaIleVal 440
 DB 1261 TGCCTATGTGTGGCTTGAACAGTGAAGATTCAAAATGAGATTCACGCTATTTGTCG 1320
 QY 441 GluPheValLeuGluLysAsnSerAsnArgAspAsnAspAspLeuProGlyLeuGlyLys 460
 DB 1321 GAGTTTGTGTTGAGATTCAGATTCGATGACATCATGATCTCTCGCATATGCAAA 1380
 QY 461 LeuLeuGluThrTyrPheLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
 DB 1381 TTGTGGAAACCTGGTTGGCGAGGGGTGTCTTCTTCTGATTCAGAGCACCAAGATGAA 1440
 QY 481 LysPheLysLeuGlyLysAspTyrTyrAspAspProMetValLeuSerTyrLeuGluVal 500
 DB 1441 AATATTAACTCGGGGACCTATGATGATGATCTTATGTTTGAAGTTACTTGAAGAGTGT 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaAlaThrMetCalaArgIleGlyAlaGlu 520
 DB 1501 GAGGTATGTCAAGGCTTCTCTTATGCTGCTGACCTATGCGCAAGATTTGAGCCGAG 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
 DB 1561 CATGTAAAGCTAGTGTATGACAGGCACTGCAAGAAATTTTCTTCCGCTATACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGlyThrValPheSerValAspProValGly 560
 DB 1621 AGAACTCGGCTGACCAAGATGTCAGAGACAGATGTTTATGTATGATTCCTGTTGCT 1680
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1681 AACCAATGTAGGCGGTATGTGAGCTGTGTCTTATGTCAGAAAGCTGTAAAGCCCTCT 1740
 QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1741 GAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGTCTTCAGAGATGAGTTGAT 1800
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluLysValLysIleLeu 620
 DB 1801 GAAACTACTCTTGAATGTCCGTTGCTGATATGTAAAGAGCAAGCAATGTGAAGATCTTA 1860
 QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSer 640
 DB 1861 GCTGCTGTGTGCAATTTGACATGATTTCACTGTTCAAGCAAGATTTTCTTAAAGC 1920
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660

DB 1921 AGCTCATCTTTTCAAGCGCAAGATATGCTTCTTATGAGAAATCTGATGTGCGACATA 1980
 QY 661 GlySerValArgAlaAspAspSerGlyAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 1981 GGGTCACTCAGACTGACATTCAGATTCAGAGCACTTCCAGAAATGTATGCTAGAGACTGACAG 2040
 QY 681 AsnIleValSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
 DB 2041 AATATGTATCCAAAGTGGCAAGAAATTAAGTCTCTGCTTTTGGCTGATACCGCAATA 2100
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAla 720
 DB 2101 GAATGTATCCAAAGGTTTGGATGCGCGAATGCTGAAGATTTGACATGACAGCAAGCT 2160
 QY 721 GluThrAlaGluLeuGlyLeuValTyrAspTyrThrLeuLeuLysSerValAspSer 740
 DB 2161 GAAACTGCGCAGCTGGGTGTGTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220
 QY 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 760
 DB 2221 GTGACAGTCTCAGACAGATGGAACCCGTGCTGTGTGAGCAACTGTGAGAGAGTGTCT 2280
 QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
 DB 2281 TGTCTATCTGATTTGGTTTCATCCGAAACCAATGCTATGTCAGAACTTACACACA 2340
 QY 781 ArgTyrGluValPheTyrSerLysSerGlyTyrLysIleThrGluGlySerValLeuAla 800
 DB 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGTGGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403

RESULT 6
 ADJ38130
 ID ADJ38130 standard; DNA; 3667 BP.
 AC
 ADJ38130;
 AC
 06-MAY-2004 (first entry)
 DE Arabidopsis thaliana Arc6-1 genomic DNA SeqID2.
 XX
 KM prokaryotic type; plasmid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; ds.
 OS Arabidopsis thaliana.
 XX
 PN W02004001003-A2.
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003W0-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Ostryoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 DR P-PSDB; ADJ38202.
 XX
 PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plasmid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX

PS Claim 1; SEQ ID NO 2; 287bp; English.

XX This invention relates to novel prokaryotic type or plastid division and CC related genes and proteins. In particular, the invention relates to novel CC Ctn2 (Arc6), Arc5 and Fzo-like genes and polypeptides. The methods and CC compositions of the present invention are useful for further CC characterizing plastid division in plant cells, in order to vary CC agronomic and horticultural characteristics of economically important CC plants, such as crop, ornamental and woody plants. They can also be used CC as herbicide targets. The present sequence is that of a DNA sequence CC which is related to the invention.

XX Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3667
Score:	3837.50	Matches:	800
Percent Similarity:	78.9%	Conservative:	0
Best Local Similarity:	78.9%	Mismatches:	1
Query Match:	94.4%	Indels:	214
DB:	12	Gaps:	5

US-10-600-070b-2 (1-801) x ADJ38130 (1-3667)

QY 1 MetG]uAlaLeuSerHisValGlyLeuSerProPheGluLeuCyArGLeuPro 20
 DB 481 ATGGAAGCTTGAGTCACGTGCGCATTTGCTCTCCCATTCACATTTATGCCATTACCA 540
 QY 21 ProAlaThrThrLyLeuArGArGSerHisAsnThrSerThrIleCySerAlaSer 40
 DB 541 CCGGCGACGACAAAGCTCCGACGTAAGCCACAACACCTCTACAACTATCTCTCGCGACG 600
 QY 41 LytTPAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 DB 601 AAATGGGCGCAGCGCTTCTCTCGACATTCATTCACCTCCGATTCCTCTCTCTCTCC 660
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 661 TTGGCCACCGCCACACACCGCCACTCTGCTCTCTGCGACCATTTATGATCGTCC 720
 QY 81 GluArgHisValProIleProIleAspPheThrGlnValLeuGlnAlaGlnThrHisPhe 100
 DB 721 GAAGCGACGTCGCCATCCCATGATTCTTACCAAGATTAGAGCTCAAAACATTTTC 780
 QY 101 LeuThrAspGlyIleArGArGAlaPheGluAlaArgAlaSerLySProProGlnPheGly 120
 DB 781 TTAAACCATGGAAATCAGAAAGCAATTCAGAGCTAGGCTTTCGAAACCGCCCAATTCGCT 840
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCySGluThrLeu 140
 DB 841 TTCAGCGACAGCGCTTTAATCAAGCCGAGACAGATTTCTCAAGCTGCTTGCAGAACTCTG 900
 QY 141 SerAsnProArGSerArGArGlyuTyArSngluGlyLeuLeuAspAspGluAlaIaThr 160
 DB 901 TCTATATCTCGCTCTAGAAAGAGTACATGAAAGTCTTCTTGATGATGAAGAGCTACA 960
 QY 161 ValIleThrAspValProTTPAsp----- 168
 DB 961 GTCATCATCTAGTCTTCTGGAGTAAGTAATTCGATTTCGAGTAATAAGTTTCTTC 1020
 QY 169 -----LySValProGlyAl 173
 DB 1021 GTTTTAAATTCATGAAATTGGATTAAGAGAACTTTTATCTAGTGAAGTTCTTGAGGC 1080
 QY 173 AluCyValLeuGlnGlnGlyGlyuThrGluIleValLeuArgValGlyGluAlaIe 193
 DB 1081 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATGTTCTTGGGGTGTGAGGCTCT 1140
 QY 193 uLeuLySGluArgLeuProLySerPheLyGlnAspValValLeuValMetAlaLeuAl 213
 DB 1141 GCTTAAGAGAGGTGTGCTTAAGTGTGTTAAGCAAGATGTGGTTTATGTTAGGCGCTTCG 1200
 QY 213 aspHeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTy 233

DB 1201 GTTTCGATGCTCCAGAGGATGCTATGCGATTGATCCACCGATTATTAATCTGTGA 1260
 QY 233 rGluPheValGluGluAlaLeuLyLeuLeuGln----- 244
 DB 1261 TGAAGTTGTGAGAAAGCTTTGAAGCTTTTACAGGATGTTTACTGCTTGTGTAATTTC 1320
 QY 244 ----- 244
 DB 1321 ACAGAGGTGGCTTTATTAAGAACTTTCTGATTGTGATTTGATTGAGCTTGTGTA 1380
 QY 245 -GluGluGlyAlaSerSerLeuAlaProAspLeuAlaGlnIleAspGluThrLeuG 264
 DB 1381 GGAGGAAGAGCAAGTAGCCTTGACACCGATTACGTGACAAATTGATGAGACTTTGGA 1440
 QY 264 uGluIleThrProArGTrValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyAl 284
 DB 1441 AGAGATCACTCCCGTTATGTCTTGAGACTTGTGCTTACCGCTTGGTGAATTAACG 1500
 QY 284 aAlaLySArGLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGly 304
 DB 1501 TCGGAAAGACTAAATGCTTTAAGCGGTGCGGAATATTTGTGCTGTTGAGAGAG 1560
 QY 304 yGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLyPheMetAsnGluAlaPheLe 324
 DB 1561 TGGAGCATGACGCTCTTGTGGGGGTTGACCCGTGAGAGTTTATGAATGAGCGCTTTT 1620
 QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
 DB 1621 ACGAATGACAGCTGTGAGAGAGT-ATPACATTTAGATACCTTTTAAATTTCTTTAGC 1679
 QY 333 -----AspLeuPheVal 336
 DB 1680 ATGATATTAATCTTAGGTTCTCATTTTAATGATGTGTGTGATGAGTTGATCTTTTTC 1739
 QY 336 lAlaThrProSerAsnIleProAlaGluSerPheGluValTyGluValAlaLeuAlaLe 356
 DB 1740 AGCTACCCCAAGCAATATTCAGACAGATTCATTTGAAGATTTTCGAAATTCACCTTCTCT 1799
 QY 356 uValAlaGlnAlaPheIleGlyLySAspProHisLeuLeuGlnAspAlaAspLySglnP 376
 DB 1800 TGTGGCTCAAGCTTTATGTTAGTAAGACACCTTTTACAGAGATGCTGATTAACGAAT 1859
 QY 376 eGlnGlnLeuGlnGlnAlaLySValMetAlaMetGluIleProAlaMetLeuTyArgP 396
 DB 1860 CCAGCACTTCACAGAGCTTAAGGTATAGCTATGAGATTTCCGCGATGTGATGATAC 1919
 QY 396 rArgAsnAsnTrpGluIleAspPheGlyLeuGluArGlyLeuCyAlaLeuLeuIleG 416
 DB 1920 ACGGAATTAATTTGGAGATAGACTTCGCTCAAGAAAGGAGCACTGTGCACTGCTTAAGG 1979
 QY 416 yLySValAspGluCyArGMetTrpLeuGlyLeuAspSerGluAspSerGlnTyArgAs 436
 DB 1980 CAAAGTTGATGATGGCGTATGTGTGGCTTACAGAGAGATTCACAAATTAAGAA 2039
 QY 436 nProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAspAspAspLeuP 456
 DB 2040 TCAAGCTATTTGAGAGTTGTTTGGAGAAITTCAAATCGATGACAAATGATCTCC 2099
 QY 456 ogLyLeuCyAlaLeuLeuGlnThrTrpLeuAlaGlyValValPheProArgPheArgAs 476
 DB 2100 TGAATATGCAAAATTGTGGAACCTGTGCGAGGGGTGTCTTCTTAAGTTTCAGAGA 2159
 QY 476 pThrLySAspLySLeuPheLySLeuGlyAspTyTyArgAspProMetValLeuSerTy 496
 DB 2160 CACCAAAAGATTAATAAATTTAACTCGGAGCTATGATGATCCATAGGTTTGAAGTA 2219
 QY 496 rLeuGluArgValGlyValValGlnLySerProLeuAlaAlaAlaIaThrMetAlaAr 516
 DB 2220 CTTGGAAGAGTGAAGATTAAGGTTCTCTTTAGCTGCTGCTGCAACTATGCGAG 2279
 QY 516 gIleGlyAlaGluHisValLySAlaSerAlaMetGlnAlaLeuGlnLySValPheProSe 536

Db	2280	GATTGAGCCGAGCATGTGAAAGCTAAGTGCATAGCAGGCACTGCAGAAAGTTTTCCTTC	233
Qy	536	TAAGTyrThrAspAlaGlnSerAlaGluProLysAspValGlnGluThrValPheSerVal	556
Db	2340	CCGCTATACAGATACAAAACCTCCGCTGAAACCCAGAGATGTGCAGAGACAGATGTTTAACTGT	2399
Qy	556	LAAspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAl	576
Db	2400	AGATCTGTGGTAAACAAATGTAGCCGAGATGTAGCGCTGTGCTTATATTCAGAAAGC	2455
Qy	576	aValArgProSerGlyAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGly	596
Db	2460	TGTAAAGACCCCTGAAAACCTTTGAAAACCTAATATATGCAANTTCAGCTGGGGCTTCAGA	2519
Qy	596	uSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSe	616
Db	2520	GAGTACGCTTGATGAAACCTACTGTTGAAATGTCCTGTCTGTATAGTTAAAGAGGCAAG	2579
Qy	616	rValLysIleLeuAlaAlaGlyValAlaIleGlyLeuLysLeuPheSerGlnLysTyr	636
Db	2580	TGTGAAGATCCTAGCTGCTGGGTGGCAATGTGACTGTATTCACCTGTACCGCAGAAAGTA	2639
Qy	636	rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs	656
Db	2640	TTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGGTTCTTCTTAAGAACTGCA	2699
Qy	656	pValAlaThrIle-----	660
Db	2700	TGTGCTACCATAGGTATGATTAATATGATGCMAATTTTCATATATCTGATTCGTCCAAAAT	2759
Qy	660	-----	660
Db	2760	ATGCTGTTTTGTGAGCTAAGAACATACTGCCACTTAATACATGTCCMAAGTTGAC	2819
Qy	660	-----	660
Db	2820	CAAGATTACAAGTTGCTGAGTAAATTTCACTAATATATGCTGCTGAAATTTTGTATCA	2879
Qy	660	-----	660
Db	2880	ACTGTAGACAGAAATGTAAATTTCACTCTCAAACTTCTGTTTGAATTAAGTAGATTA	2939
Qy	660	-----	660
Db	2940	GAGATTCGCTTAGTGTGCTTGTTCCACTTTCTTTCCTGATTTTTCCTTTTCGATT	2999
Qy	661	--GlySerValArgAlaAspAspSerGlyAlaLeuProArgMetAspAlaArgThrAlaG	680
Db	3000	TAGGCTACGTCAAGCTCACAATTCAGAACGACCTTCCAGATAGATGCTTAGACTGCAG	3059
Qy	680	lAsnIleValSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHisArgI	700
Db	3060	AGAAATATATGATCCAAAGTGGCAAGATTAAATGTCCTGGCTTTTGGGCTGATCACCGCA	3119
Qy	700	lGluMetLeuProGlu-----	705
Db	3120	TAGAAATGTACCAAGCGTGAAGGCAATTAATCTCAATTCATCAATCAATTTGTGAAAAC	3179
Qy	706	-----ValLeuAspGly	709
Db	3180	TTGGACATGATTATAGTCTGCTGCTGCTGTTGATCTGTTATTTATTAAGTTTTTGGATGGG	3239
Qy	710	ArgMetLeuLysIleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyr	729
Db	3240	CGAATGCTGAAGATTTTGACATGCAGACAGCACTGAACCTGGCAGCTTGGGTTGGTTAT	3299
Qy	730	AspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArg	749
Db	3300	GATATATACACTGTGAACATATCTGTTGACAGTGCAGACTCAGCAGATGAACCCGT	3359
Qy	750	AlaLeuValGluAlaThrLeuGlnGlnSerAlaCysLeuSerAspLeuValHisProGlu	769
Db	3360	GCTCTGGGTGAAGCACTGTGAAGAGCTGTGCTGTCTATCTGATTTGGTTCAATCCAGAA	3419

Qy	770	AenlnsalatlrAspValaArgThrLythrhThraglyGluValPheTrpSerIysSer	789
Db	3420	AACATGCTACTGATGTCAGAACCTACACACAAACATACGAACTTTTCGTCTCAAGTCA	347
Qy	790	GLYTrpIysIleThrGluGlySerValIleuIaSer	801
Db	3480	GGGTGGAAATCATCAGAAAGGCTCTGCTTTCGATCA	3515

RESULT 7

ID ADJ38136 standard; DNA; 3667 BP.

AC ADJ38136;

DT 06-MAY-2004 (first entry)

DE Arabidopsis thaliana AtFtn2 genomic DNA SeqID10.

KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;

KW herbicide target; ds.

05 *Arabidopsis thaliana*.

PN WO2004001003-A2

PD 31-DEC-2003.

PF 20-JUN-2003; 2003WO-US019536.

PR 20-JUN-2002; 2002US-0390140P.

PR 20-JUN-2003; 2003US-00600070.

PA (UNMS) UNIV MICHIGAN STATE.

PI Osteryoung KM, Vittha S, Koksharova OA, Gao H;

DR WPI; 2004-082486/08.

X

PT For further characterizing plastid division in plant cells, and in

PT important plants.

PS Claim 1; SEQ ID NO 10; 287pp; English

CC This invention relates to novel prokaryotic type or plastid division and

CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and

CC characterising plastid division in plant cells, in order to vary

CC plants, such as crop, ornamental and woody plants. They can also be used

CC which is related to the invention.

SQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Alignment Scores:

Score: 3823.50

Best Local Similarity: 78.7%

DB: 12

US-10-600-070B-2 (1-801) X ADJ38136 (1-3667)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCybArgLeuPro 20

Db 481 ATGGAAGCTCTGAGTCACGTCGGCATTTGGTCTCTCCCATTCACATTATGCCGATTACCA 540
 QY 21 ProLaThThrThyIeLeuArgArgSerH1eAsnThSerThrThr1IeCysSerAlaSer 40
 Db 541 CCGGCACACAAAGCTCCGACGTAGCCACACACCTCTACACTATCTGCTCCGACAG 600
 QY 41 LysTrpAlaAspArgLeuSerSerAspPheAsnPheThSerAspSerSerSerSer 60
 Db 601 AATATGGCCGACCGCTCTTCTCTCCGACTTCATTTCACTCCGATTCCTCTCTCTCC 660
 QY 61 PheAlaThAlaThThrThrThrAlaThrLeuValSerLeuProProSer1IeAspArgPro 80
 Db 661 TTCGCCACCGCCACACACCGCACCTCGTCTCTCTGCCACCATCTATTTGATTCCTCC 720
 QY 81 GluArgHisValaPro1IePro1IeAspPheTyArg1ValaLeuGly1aGlnThrH1ePhe 100
 Db 721 GAACGCCACGTCGCCATCCCATTTGCTTACACAGATTTAGACCTCAACACACTTC 780
 QY 101 LeuThrAspGly1IeArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
 Db 781 TTAAACGATGGATTCAGAAAGACATTCGAAAGCTTAGGTTTCAAAACCGCCGAACTCGGT 840
 QY 121 PheSerAspAspAlaLeu1IeSerArgArgGln1IeLeuGln1aAlaCysGlnThrLeu 140
 Db 841 TTCAGGAGACGCTTTAATCAGCCGAGACAGATTCCTCAAGCTGCTTCCGAAACTCTG 900
 QY 141 SerAsnProArgSerArgArgGlyTyArgGlnGly1LeuLeuAspArgGlyGluAlaThr 160
 Db 901 TCTAACTCTCGCTCTAGAAAGAGTACMAATGAAGGCTCTTGATGATGAAGAGCTAC 960
 QY 161 Val1IeThrAspValaProTrpAsp----- 168
 Db 961 GTATCATCGATGTCTTGGAATAGGTAAAGTATTCGATTCGGAATATTAAGTTCTTC 1020
 QY 169 -----LysValaProGlyAla 173
 Db 1021 GTTTTAATTTCAAGAAATGGAATGAAGAAAGAACTTTATCTAGTAAAGTCTCGGGGC 1080
 QY 173 AleuCysValaLeuGlnGlnGlyGlyGluThrGln1IeValaLeuArgValaGlyGluAla 193
 Db 1081 TCTCTGTATCTCAAGAAAGTGTGAGACTGAGATAGTTCTTCGGGTGTGTAAGGCTCT 1140
 QY 193 UleuLysGluArgLeuProLysSerPheLysGlnAspValaValaLeuValaMetAlaLeuAla 213
 Db 1141 GCTTAAAGAGAGGTGTCTAAGTCTTAAAGCAAGATGTGTTTACTTAAAGGCTGTG 1200
 QY 213 aPheLeuAspValaSerArgAspAlaMetAlaLeuAspProProAspPhe1IeThrGlyTy 233
 Db 1201 GTTTCGATGTCTCGAGGGATGTCTAAGCATTTGCATCCACTGATTTTATTTACTGGT 1260
 QY 233 rGluPheValaGlnGlnAlaLeuLysLeuLeuGln----- 244
 Db 1261 TGAAGTTTGTGAAGAACTTTGAAGCTTTTACAGGTAGTTGACTTCCTTGGTAATTTG 1320
 QY 244 ----- 244
 Db 1321 ACAGCGCTGGCTTTATTAAGAACTTTGATTTGATTAAGTCTTGTGTA 1380
 QY 245 -GlnGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGln1IeAspGlnThrLeuGln 264
 Db 1381 GGAAGAAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTAAGAGCTTTGGA 1440
 QY 264 uGlu1IeThrProArgTyValaLeuGlnLeuLeuGly1LeuProLeuGln1IeAspArgTyAla 284
 Db 1441 AGAGATCACTCCGCTTATGTCTTGAAGCTAAGCTTACCGCTTACCGTGTGATTAAGC 1500
 QY 284 aAlaLysArgLeuAsnGlyLeuSerGlyValaArgAsn1IeLeuTrpSerValaGlyGly 304
 Db 1501 TGGGAAAAAGACTAATGTGTTAAGCGGTGTGCGGAAATATTTGTGTGTGAGAGAG 1560
 QY 304 yGlyAlaSerAlaLeuValaGlyGlyLeuThrArgGlyLysPheMetAsnGlnAlaPheLe 324
 Db 1561 TGGAGCATCAAGCTCTTGTGGGGGTTTGAACCGGTGAAGAGTTTATGAATAGGCGTTT 1620

QY 324 uArgMetThralaAlaGlnGlnVala----- 332
 Db 1621 ATGAATGACAGCTGTGAGACAGT-ATACAGTTAGATACCTTTTATTTCTTTAGC 1679
 QY 333 -----AspLeuPheVala 336
 Db 1680 ATGATTAATCTTAAAGTTTCATTTTAATGATGTGTGTGTGAGGTTAGCTTTTGT 1739
 QY 336 1AlaThrProSerAsn1IeProAlaGlySerPheGluValaTyArgValaAlaLeuAlaLe 356
 Db 1740 AGCTACCCACAGCAATATTTCCAGAGATCATTTTAAGTTTACGAAGTTGCACCTGCTCT 1799
 QY 356 uValaAlaGlnAlaPhe1IeGlyLysLysProH1eLeuLeuGlnAspAlaAspLysGlnH 376
 Db 1800 TGTGCTCAAGCTTTTATTTGTTAAGGACCACTTTTACAGGATGTGATTAAGCAATT 1859
 QY 376 eGlnGlnLeuGlnGlnAlaLysValaMetAlaMetGlu1IeProAlaMetLeuTyTrpH 396
 Db 1860 CCAGCAACTTCAGCAGGCTTAAGGTAAAGCTATGAGATTCCTCGCATGTGTATGATAC 1919
 QY 396 rArgAsnAsnTrpGlu1IeAspPheGlyLeuGlnArgGlyLeuCysAlaLeuLeu1IeG 416
 Db 1920 ACGGATTAATGTGAGATAGACTCGGTCAAGAAAGGAACTCTGTGCACCTGCTTAAG 1979
 QY 416 yLysValaAspGluCysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyArgAs 436
 Db 1980 CAAGGTGATGAATGCCATGTGTGCTTGAAGCACTGAGCAATTCACAAATATAGAA 2039
 QY 436 nProAla1IeValaGluPheValaLeuGlnAsnSerAsnArgAspAspAspLeuPr 456
 Db 2040 TCCAGCTATTGTGAGTTGTGTTTGAGAAATCAATCTGATGACMAATGATGATCTCC 2099
 QY 456 oGlyLeuCysLysLeuLeuGlnThrThrLeuAlaGlyVala1IePheProArgPheArgAs 476
 Db 2100 TGGACTATCAATATGTTGAAACCTGTGTGACAGGGGTGTCTTCTTCAAGTTCAAGA 2159
 QY 476 rPheLysAspLysLysPheLysLeuGln1IeAspTyTyTrpAspAspProMetValaLeuSerTy 496
 Db 2160 CACCAAAAGATTAATAATTTAACTCGGGGACTACATATGATCCATGTGTTTGATTA 2219
 QY 496 rLeuGlnArgValaGluValaValaGln1IeSerProLeuAlaAlaAlaThrMetAlaAr 516
 Db 2220 CTTGGAAAGGTGAGAGTATGTCAGGGTCTCTTATAGCTGCTGACGACTATGGCAAG 2279
 QY 516 g1IeGlyAlaGlnHisValaLysAlaSerAlaMetGlnAlaLeuGlnLysValaPheProSe 536
 Db 2280 GATTGAGCGGACGATGTAAAGCTAGTCTATGACGGACTGCAAGAAAGTTTCTTTC 2339
 QY 536 rArgTyTrpAspArgAsnSerAlaGluProLysAspValaGlnLysThrValaPheSerTy 556
 Db 2340 CCGCTATACAGATGAATACTCGGCTGAACCAAGATGTGCAAGAAACAGTGTATGTGT 2399
 QY 556 1AspProValaGlyAsnAsnValaGlyArgAspGlyGluProGlyValaPhe1IeAlaGluAl 576
 Db 2400 AGATCTGTGTGTTGTTAACAATGATGAGCGGTATGTGACCTGTGTATTTGCAAGAC 2459
 QY 576 aValaArgProSerGluAsnPheGluThrAsnAspTyTral1IeArgAlaGlyValaSerG 596
 Db 2460 TGTAAACCTCTGAATAAATTGAAACTAATGATTTTGCAATTCAGCTGGGGCTTCAGA 2519
 QY 596 uSerSerValaAspGluThrThrValaGlnMetSerVala1IeAspMetLeuLysGln1IeSe 616
 Db 2520 GAGTACGTTGATGAACCTATCTTAAATGTCCTGTGCTGATATGTTAAAGAGCCAG 2579
 QY 616 rValLys1IeLeuAlaAlaGlyVala1IeGlyLeu1IeSerLeuPheSerGlnTyTy 636
 Db 2580 TGTGAAGATCTCTAGCTGTGTGTGCAATTTGACATTTTCACTGTTCAGCCAGAAAGTA 2639
 QY 636 rPheLeuLysSerSerSerPheGlnArgLysAspMetValaSerSerMetLysSerAs 656
 Db 2640 TTTCTTTAAAGACGCTCATCTTTTCAACGCAAGGATATGTGTTCTTCTATGGAATCTGA 2699

QY 656 PValAlaThrIle----- 660
 DB 2700 TGTGGCTACCATGATGATTAATGATGCAATTTTCATATATCTGCATTCGCAAAAT 2759
 QY 660 ----- 660
 DB 2760 ATGCTGTGTTTGAGCTAAGAACATAGTCCCACTTAATACATGTCCCAAAAGTTGTAC 2819
 QY 660 ----- 660
 DB 2820 CAAGATTACAGTGTCTGAGTAATTTCACTAATATAGCTGTGAATTTTGTATCA 2879
 QY 660 ----- 660
 DB 2880 ACTGTAGACAGAAATGTAATTTCACTCTCACTCAATTTCTGTATGAATAAGATGAATTA 2939
 QY 660 ----- 660
 DB 2940 GAGATTGCTTACGTGTGCTTTGTCCAACTTTCTTCTGTGATTTTCTTTTCGATT 2999
 QY 661 --GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
 DB 3000 TAGGGTCAGTCAAGAGCTGAGATTCAGAACCACTTCCAGAAATGATGCTAGAGACTGCGAG 3059
 QY 680 1uAnu11ValSerIleValTrpGlnIleValIleValSerLeuAlaPheGlyProAspHisArgI 700
 DB 3060 AGAATATAGTATCCAAATGCGACAGAAATTAAGTCTGCGCTTTGGGCTGATCACCGCA 3119
 QY 700 1eGluMetLeuProGlu----- 705
 DB 3120 TAGAAATGTTACAGAGGTGAGGGAATTAATCTACATTCATCAATTTGTGMAAACTG 3179
 QY 706 -----ValIleAspGly 709
 DB 3180 TTGGACATGATTATAGTCTGTGCTGCTTTGATTGATTCTGTTATTAAGTTTGGAGCGG 3239
 QY 710 ArgMetLeuValIleTrpThrArgAspAlaAlaGluThrAlaGlnLeuGlyLeuValIle 729
 DB 3240 CGAATGCTGAAGTTTGGATGTCAGACAGAGCTGMAACTGCGAGCTTGGGTTGGTTTAT 3299
 QY 730 AspTyThrIleLeuValSerValAspSerValIleAspSerValThrValSerAlaAspGlyThrArg 749
 DB 3300 GATTATACACTGTTGAACATCTATCTGTGACAGTGTACAGTCTCAGACAGATGGAACCCGT 3359
 QY 750 AlaIleValGluAlaThrIleGlnGluSerAlaCysLeuSerAspLeuValHisProGlu 769
 DB 3360 GCTCTGTGAGAACCACTCTGAGAGAGTGTCTGTATCTGATTTGTTCAATCCAGAA 3419
 QY 770 AsnAspAlaThrAspValArgThrTyThrArgTyrGluValPheTrpSerIleValSer 789
 DB 3420 AACAAATGCTACTGATGTCAGAACCTTACACAAAGATACGAAGTTTCTGGTCCAAAGTCA 3479
 QY 790 GlyTrpIleIleThrGlnGlySerValLeuAlaSer 801
 DB 3480 GGGTGAATAATCACTGAAGGCTCTGTTCTTGATCA 3515
 RESULT 8
 ADJ38206 standard; DNA; 2283 BP.
 AC ADJ38206;
 DT 06-MAY-2004 (first entry)
 DE Placid division-related Arc6 orthologue gene 1.
 KW prokaryotic type; placid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; gene; ds.
 OS Oryza sativa.
 PN MO2004001003-AZ.

XX 31-DEC-2003.
 PD 20-JUN-2003; 2003MO-US019536.
 XX 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX (UNMS) UNIV MICHIGAN STATE.
 PA Oosteryoung KM, Vilha S, Koksharova OA, Gao H;
 PI WPI; 2004-082486/08.
 DR P-PSDB; ADJ38205.
 XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing placid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX Disclosure, Fig 8, 287pp, English.
 XX This invention relates to novel prokaryotic type or placid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing placid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.
 XX SQ Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2, 6e-154 Length: 2283
 Score: 1775.50 Matches: 390
 Percent Similarity: 62.5% Conservative: 119
 Best Local Similarity: 47.9% Mismatches: 213
 Query Match: 43.7% Indels: 93
 DB: 12 Gaps: 16
 US-10-600-070b-2 (1-801) x ADJ38206 (1-2283)
 QY 12 SerProPheGlnLeuValArgLeuProAlaThrThrIleValArgSerHisAsn 31
 DB 40 GCGCATTCGCGCTTCTCCCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCT 99
 QY 32 ThrSerThrThrIleCys---SerAlaSerIleValAlaAspArgLeuLeuSerAspPhe 50
 DB 100 CACCCCTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGCGACTTC 159
 QY 51 AsnPhe-----ThrSerAspSerSerSerSerSerPheAlaThrAlaThr 65
 DB 160 CACCTCTCCCAACCG 219
 QY 66 ThrThrAlaThrIleValSerLeuProProSerIleAspArgProGluArgHisValPro 85
 DB 220 GCGCCCTCCGCGCTCCCTTCGCGCGCTTCCTCCCGAGCGCGCGCGCGCGCGCGCGCG 279
 QY 86 IleProIleAspPheTyrglnValIleGlnValAlaGlnThrHisPheLeuThrArgGlyIle 105
 DB 280 CTCCAAGTGAATTTCTTCAAGGTTCTTGAAGGAGCAACCAATTTCTTGGCGATGGCATTC 339
 QY 106 ArgArgAlaPheGlnAlaArgValSerIleValProProGlnIlePheGlyPheSerAspAla 125
 DB 340 AGAAGGCGCTTTCAGAGCAGAGTACGACCAACCGCATATAGTCTACAGCAGCATGCT 399
 QY 126 LeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrIleSerAsnProArgSer 145
 DB 400 CTTGTTGTCGTCGACAAATGCTGACAGATTGCCATGACACTCTCATGAACCAAGAACTCC 459

Db 1 GCAGTTGACATGCTGGGGNGAATTCACAGTGAATTTACAGACGAGCCTTCTTG 60
Qy 325 ArgMetThrAlaIaIaGluGlnValAspLeuPheValAlaThrProSerAsnIleProAla 344
Db 61 CATATGACTGCACTGAGGAGTTGATTTATTTGACTGATCCCACTAATATATCCGGCA 120
Qy 345 GluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLys 364
Db 121 GAAGAGTTTGAAGTTTATGCGGTGCTGCTTCCGCTTCTTCTCAAGCTTTGTTGTA 180
Qy 365 LysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnGlnAlaLysVal 384
Db 181 AAACCTCATCATCTCAATGATGCTGAAGAACCTATTCAGAAACCTTCAGCAGCTTAAGTA 240
Qy 385 MetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIleAspPhe 404
Db 241 ACAGCTGTAGACATTTCTTTCACACTAATATACCAAGAAAGCACTGATAGACTTT 300
Qy 405 GlyLeuGluArgGlyLeuCybAlaLeuLeuIleGlyLysValAspGluCybArgMetTrp 424
Db 301 GCTTTGAGAGGGAGCTCTGTTCACTTCTTCAAGGAGCCTTGATGACAGTCTGCTGG 360
Qy 425 LeuGlyLeuAspSerGlnAspSerGlnTyrArgAsnProAlaIleValGluPheValLeu 444
Db 361 TTGGGCTTACAGCATATGATTCACATATAGAAATCATCTGTGTGATCTTTGCTTG 420
Qy 445 GluAsnSer---AsnArgAspAspAspAsp-----AspLeuProGlyLeuCybLys 460
Db 421 GAGAACTCAAGATGACATGACAAATGACAAATGACAAATGACAAATGACAAATGACAA 480
Qy 461 LeuLeuGluIleThrTyrPheValAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
Db 481 CTATTGAGAGCGTGTGTTGATGAGAGTGTATTCCTCCAGCTTTAGAGACACCAAGACATA 540
Qy 481 LysPheLeuLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGluVal 500
Db 541 GAGTTGACATGCGGAGCTACTATGATGATCTTCACTTGAATCTTGAATCTTGAATCTTGA 600
Qy 501 GluValValGlnGlySerProLeuAlaAla 510
Db 601 GATGCGACATATGTTTCACTTACCTTACCTGCT 630

RESULT 11
ADJ38215
ID ADJ38215 standard; cDNA; 660 BP.
XX
AC ADJ38215;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plasmid division-related Arc6 orthologue cDNA 8.
XX
KW prokaryotic type; plasmid division; Fun2; ARCS; ARCS; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
OS Medicago truncatula.
XX
OS
XX
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
XX
PR 09-AUG-2002; 2002US-040242P.
XX
PR 20-JUN-2003; 2003US-00600670.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oesteryoung KW, Vitcha S, Koksharova OA, Gao H;
XX
XX
DR WPI; 2004-082486/08.

XX
PT New isolated Fun2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plasmid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX
PS Disclosure; Fig 8; 287pp; English.
XX
CC This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fun2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6.5e-57 Length: 660
Score: 718.00 Matches: 141
Percent Similarity: 79.5% Conservative: 35
Best Local Similarity: 64.1% Mismatches: 34
Query Match: 17.7% Indels: 10
DB: Gaps: 4
US-10-600-070B-2 (1-801) x ADJ38215 (1-660)
Qy 25 LysLeuArgArgSerHisAsnThrSerThrIleCybSerAlaSerLysTrpAlaAsp 44
Db 15 AAACCTAACCGTCTCAT-----TCTCCGCGCTCCGCCACCATTAATGCGGAG 68
Qy 45 ArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSerPheAlaThrAla 64
Db 69 CGACTCATTTCCGATTTCCCAATTCCTCGGCGACACTCTCTTCCCTCCACCAACACC 128
Qy 65 ThrThrThrAlaThrLeuVal---SerLeuProProSerLysAspArgProGluArgHis 83
Db 129 TCCGCCACAGTCACTCTCACTCTTCTTACCTCTCCCAATA-----GAACCCAC 179
Qy 84 ValProIleProIleAspPheTyrGlnValLeuGlyValAlaGlnThrHisPheLeuThrAsp 103
Db 180 GTGTCACTCCCTCTCGACCTGTCATCAAAATCTCGGCGGAACGATTTCTCGGTGAT 239
Qy 104 GlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGlyPheSerAsp 123
Db 240 GGTATTGAGAGAGCTTATGAAAGCAAAATCTCGAACCTCTCAGATGCTTCACTAAT 299
Qy 124 AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCybGluThrLeuSerAsnPro 143
Db 300 GAAGCTTTGATTAAGTGTGCTGTCAGATCTTCAAGCTTGTGAACCTTACGTATCT 359
Qy 144 ArgSerArgArgGluTyrAsnGlnGlyLeuLeuAsp-----AspGluGluAla 159
Db 360 GCTTCTGAAGAGAGATATATCAAAAGCTCTGTGAGATGAAGAGCAAGATGCAATCT 419
Qy 160 ThrValIleThrAspValProTrpAspLysValProGlyAlaLeuCybValLeuGlnGlu 179
Db 420 TCCATTTCTCACTGAATCTCTTTCGACAAAGTTCTGTGAGCTCTGTGCGGTGCAAGAA 479
Qy 180 GlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuArgGluTrpLeuPro 199
Db 480 GCTGAGAGACGAGAGTTGGTCTTGGATTTGAGGGGCTTTATGAGAGAGGTTACCG 539
Qy 200 LysSerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArg 219
Db 540 AAGATGTTTAAGCAAAATGTGTGTGCTGCTATGCGCTTGCATATGTTGACGTTTCTAG 539
Qy 220 AspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAla 239
Db 600 GATGCTATGGCTTTGTCCCGCCAGATTTCAATGTTGCTTGTGAGATGCTGGAAGGCA 659

RESULT 12
ADJ38226
ID ADJ38226 standard; cDNA; 622 BP.
XX
AC ADJ38226;
XX
DT 06-MAY-2004 (first entry)
XX
DE Placoid division-related Arc6 orthologue cDNA 19.
XX
KM prokaryotic type; placid division; Ftn2; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ss.
OS Hordeum vulgare.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oteyroung KW, Vitsha S, Kokeharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing placid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or placid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing placid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 622 BP; 163 A; 121 C; 155 G; 178 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 5,53e-41 Length: 622
Score: 546.00 Matches: 117
Percent Similarity: 69.3% Conservative: 30
Best Local Similarity: 55.2% Mismatches: 50
Query Match: 13.4% Indels: 16
DB: Gaps: 5

US-10-600-070b-2 (1-801) x ADJ38226 (1-622)

QY 266 ILeThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyrAlaAla 285
DB 17 ATCAACCCCTCGTGTGTTTAAAGACTTCTTGCCCTCTCTT-----GACGAGNAAACA 70
QY 286 LybArgLeuLen-----GlyLeuSerGlyValArgLeuLeuTyrSerValGly 302
DB 71 CC-AGAGTAAACGCCAAGNAAAGCTTCCTGCTGTGAGAAACATTTTGTGAGGTGGT 129
QY 303 GlyGlyValAspSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAla 322
PS 130 AGAGAGGTATTTCTACTGTTGTGAGAGATTTTACGAGAGCCTACATGAATGAGGCC 189

QY 323 PheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSerAsnIle 342
DB 190 TTTTGGAGATGACATGACTGACAGATGATGATTTCTTTCAAAACGCCGAATGACATA 249
QY 343 ProAlaGluSerPheGluValTyrGluValAlaLeuValaLeuValaGlnAlaPheIle 362
DB 250 CCACCTGAATGTTGTAATCTATAGCGTGGCACTGCCAATGTTGCTCAAGCAATTGTA 309
QY 363 GlyValLysPProHleuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnAla 382
DB 310 AGTAAAGGCCAGAGCTCATCATGTCGAGATGATCTTTTCGAACGCTCCAGAAATTC 369
QY 383 LybValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAspAsnTyrGluIle 402
DB 370 AATATCGCTTCTCA-----TATCCTTATGCTAAC-----GAGATG 405
QY 403 AspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyLysValaAspGluCyAla 422
DB 406 GATCTTGCGTTGGAAAGGCACTTGTCTCATTTGTTGGAGACATTAGCACTGCAGA 465
QY 423 MetTyrLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPhe 442
DB 466 ACTTGCGCTTCGATTGATTAATGATCTTACACATAGAGACCCGAATAATTGTAGAGTTT 525
QY 443 ValLeuGluLysSerAsnArgAsp-----AspAsnAspAspLeuProGlyLeuCyAla 460
DB 526 ATGTGAACAACCTTACATTCGACACACAGAGAGATGATCTTCCAGGCGCTGTGTAG 585
QY 461 LeuLeuGluThrTyrTrpLeuAlaGlyValaIlePhePro 472
DB 586 CTTTGGAGACTTGCGCTTGTCTCAGAGGTTTTCCT 621

RESULT 13
ADJ38223
ID ADJ38223 standard; cDNA; 537 BP.
XX
AC ADJ38223;
XX
DT 06-MAY-2004 (first entry)
XX
DE Placid division-related Arc6 orthologue cDNA 16.
XX
KM prokaryotic type; placid division; Ftn2; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ss.
XX
OS Triticum aestivum.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oteyroung KW, Vitsha S, Kokeharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing placid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or placid division and

CC related genes and proteins. In particular, the invention relates to novel
 CC Fm2 (AHC6), AHC5 and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.

XX
 SQ Sequence 537 BP; 133 A; 113 C; 151 G; 140 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,31e-40	Length:	537
Score:	541.00	Matches:	111
Percent Similarity:	76.3%	Conservative:	24
Best Local Similarity:	62.7%	Mismatches:	42
Query Match:	13.3%	Indels:	0
DB:	12	Gaps:	0

US-10-600-070b-2 (1-801) x ADJ38223 (1-537)

QY 186 ValLeuAryValGlyGluAlaLeuLeuLysGluArgLeuProLysSerPheLysGlnAsp 205
 DB 3 GTGCTTCGCAATTCGAGGCGCACTTACGAGACCGCCGCAAGCGGCTTCAGACAGAT 62
 QY 206 ValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAsp 225
 DB 63 GTGCTTCGCAATTCGAGGCGCTTCGCTTATGATCTTACAGAGACGAGATCGCGCTTACG 122
 QY 226 ProProAspPheLeuThrGlyTyrGluPheValGluGluAlaLeuLysLeuLeuGlnGlu 245
 DB 123 CTCCTCGATGTAATCCGCTCTGAGGTGCTTGAAGGCGCTCTCAAGCTTTTGCAAGAG 182
 QY 246 GluGluValAserSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGluGlu 265
 DB 183 GATGGGGCAATCAATTCGACCTGCTGCTCCTCAACAAATGATGAACTCTGAGAGAT 242
 QY 266 IleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspArgTyrAlaAla 285
 DB 243 ATCACTCTGCTGCTGCTTTCGAGCTTCCTCTCTTATGAAAAACATCAGAT 302
 QY 286 LysArgLeuAnGlyLeuSerGlyValArgAsnIleLeuTyrSerValGlyGlyGly 305
 DB 303 GAACACCAAGAGGCTTCGCGGTGAGAAACATTTTGAGAGTGTGGCAGAGAGGT 362
 QY 306 AlSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetLeuGluAlaPheLeuArg 325
 DB 363 ATGTGTACTGTTGAGAGGATTTTCCGCGAAGCCCTACATGATGAAAGCCTTCCTGACG 422
 QY 326 MetThrAlaIleGluGluValAspLeuPheValAlaThrProSerAsnIleProAlaGlu 345
 DB 423 ATGACATCGCGGAGGAGATGATTTTCTTCAAAAACCGAATACATACCGCTCGAA 482
 QY 346 SerPheGluValTyrGluValAlaLeuAlaLeuValaGluAlaPheIle 362
 DB 483 TGGTTGAATCTATACGTGACCTTGCAAAATGTTCTCAAGCAATTTGA 533

RESULT 14
 ACN48781
 ID ACN48781 standard; cDNA; 607 BP.
 XX
 AC ACN48781;
 DT 02-DEC-2004 (first entry)
 XX
 XX Cotton primed seed EST Clone ID: LIB3825-027-06-K6-H1, SEQ:3562.
 XX
 KW Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
 KW variety DP50B; library LIB3825; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.
 XX
 XX Gosypium hirsutum.

XX
 PN US2004123340-A1.
 XX
 XX 24-JUN-2004.
 PD
 XX 12-DEC-2001; 2001US-00021323.
 PF
 XX 14-DEC-2000; 2000US-0255619P.
 PR
 XX
 XX (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K J.
 PA (ZIEG/) ZIEGLER T E.
 XX
 XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 WPI; 2004-479808/45.
 DR
 XX
 XX
 PT New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 PT
 XX
 XX Claim 1, SEQ ID NO 3562; 34pp; English.
 PS
 XX
 XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nicotiana glauca. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?DocID=US20040123340
 XX
 SQ Sequence 607 BP; 172 A; 103 C; 164 G; 167 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 2.02e-39 Length: 607
 Score: 529.00 Matches: 120
 Percent Similarity: 64.0% Conservative: 22
 Best Local Similarity: 54.1% Mismatches: 44
 Query Match: 13.0% Indels: 36
 DB: 13 Gaps: 5

US-10-600-070b-2 (1-801) x ACN48781 (1-607)

QY 420 GluCyArgMetThrLeuGlyLeuAspSerGluAspSerGlnTyrArgAspProAlaIle 439
 DB 9 GAATGCGGCTTCGCTGAGGCTTACAGAGATGAGCTCCCTTATGAAATATACATCTATT 68
 QY 440 ValGluPheValLeuGluLeuSerAsnArgAspAspAsnAspLeuProGlyLeuCy 459


```

Db      69  GTAGAAATTGTCCTGGAAGAACTCAAGAGATGACAGATCTCCCGGCTTGC 128
Qy      460  LyeleuleuGluThrTTTPLeuAlaGlyValAlaPheProArgPheArgPheThylAsp 479
Db      129  AAATCTGTGAGGAGGATGCTAATGAGAGTGGTTTCTTCAATTAGACACCAAGAT 188
Qy      480  LyeleuPheLysleuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGluArg 499
Db      189  ATCAATTCAGACTTGAGATTTATGATGATCTTACTGCTCGAGATATTTGAAAGG 248
Qy      500  ValGluValValGlnGlySerProLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 519
Db      249  CTGAGGAGAGAGGTGTTCACTTGGCTGACCCGAGCTATAGAGAGATAGGTGCT 308
Qy      520  Glu-----HisValLysAlaSerAlaMetGlnAlaLeuGluLysVal 533
Db      309  GAGGCTACTGCACTTCTTGTATCATGTAAAGCTAGTCAATTGAGCATTCGAGAGGTG 368
Qy      534  PheProSerArgTyrThrAspArgAsnSerAlaGluProLysAspValGlnGluThrVal 553
Db      369  TTTCCTCTTGCTGCTGCAAGAGAGACTGCGAGACATCAA----- 407
Qy      554  PheSerValAspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValAlaPhe 573
Db      408  -----TTAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 434
Qy      574  AlaGluAlaValAlaArgProSerGluAsnPheGluThr-----AsnAsp 587
Db      435  -----CCTGTTGAAAGTGAAGAGACTCTTGGAACCTGATCAAGAGAGAT 479
Qy      588  TyrAlaIleArgAlaGlyVal-----SerGluSerSerValAspGluThrThrValGluMet 606
Db      480  TCTGCATCTCTAGCTGAGGTTCTCGAATTAAGCACTTGGAGAGATGATTAAGAGAA 539
Qy      607  SerValAlaAspMetLeuLysGluAlaSerValLysIleLeuAlaAlaGlyValAlaIle 626
Db      540  ACATATAGTGAACAATAATTAAGATGCCAGTCTCAAGATCATGCTGCTGTGTGTAAAT 599
Qy      627  GlyLeu 628
Db      600  GGAGTG 605

RESULT 15
ADJ38246
ID      ADJ38246 standard; cDNA; 2307 BP.
XX      AC      ADJ38246;
XX      DT      06-MAY-2004 (first entry)
XX      DE      Placoid division-related Arc6 orthologue cDNA 35.
XX      KM      prokaryotic type; placoid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
XX      KM      agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX      KM      herbicide target; gene; ss.
XX      OS      unidentified.
XX      PN      MO2004001003-A2.
XX      PD      31-DEC-2003.
XX      PF      20-JUN-2003; 2003WO-US019536.
XX      PR      20-JUN-2002; 2002US-0390140P.
XX      PR      09-AUG-2002; 2002US-0402242P.
XX      PR      20-JUN-2003; 2003US-00600070.
XX      PA      (UNMS ) UNIV MICHIGAN STATE.
XX      PI      Oesteryoung KM, Vitcha S, Kokeharova OA, Gao H;
XX

```

```

DR      WPI; 2004-082486/08.
DR      P-PSDB; ADJ38247.
XX      PT      New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
XX      PT      for further characterizing placoid division in plant cells, and in
XX      PT      varying agronomic and horticultural characteristics of economically
XX      PT      important plants.
XX      PS      Disclosure; Fig 8; 287pp; English.
XX      CC      This invention relates to novel prokaryotic type or placoid division and
XX      CC      related genes and proteins. In particular, the invention relates to novel
XX      CC      Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
XX      CC      compositions of the present invention are useful for further
XX      CC      characterizing placoid division in plant cells, in order to vary
XX      CC      agronomic and horticultural characteristics of economically important
XX      CC      plants, such as crop, ornamental and woody plants. They can also be used
XX      CC      as herbicide targets. The present sequence is a cDNA sequence which is
XX      CC      related to the invention.
XX      SO      Sequence 2307 BP; 724 A; 490 C; 480 G; 613 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      3,56e-38      Length:      2307
Score:          524.50      Matches:      212
Percent Similarity: 39.9%      Conservative: 128
Best Local Similarity: 24.9%      Mismatches: 286
Query Match:    12.9%      Indels:      227
DB:             12      Gaps:        28

US-10-600-070B-2 (1-801) x ADJ38246 (1-2307)
Qy      84  ValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPheLeuThrAsp 103
Db      1  GTGGAGATTCCTGCTGATTAACCACTTACCGAATTTAGACTACCGTTAGCGCAAGTGAAGAA 60
Qy      104  GlyIleArgAlaPheGluAlaArgValSerLysProGlnPheGlyPheSerAsp 123
Db      61  CAATTGCGACAGGAGATACAGGATCGATTTGACATTTGCCACGACCTGATATTTCCAG 120
Qy      124  AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrSerAspPro 143
Db      121  GCAGCAATTTCTTCTGTAACAACATCATAGAGAACTTACGTTTATCATGATCA 180
Qy      144  ArgSerArgArgGluTyrAsnGluGlyLeu-----AspAspGlu 157
Db      181  AAACACAGCAGATCACTACGATCAGCTTATCTTGCCCGCTATGACCTGTAAACCTT 240
Qy      158  GluAlaThrValIleThr----- 163
Db      241  GCTGCTGCGCAGTAGACAGAAATCGTACAGAAAGCACCAAAAGGCGTAGATACC 300
Qy      164  -----AspValProTyrAspLysValProGlyAlaLeuCysValLeuGln 178
Db      301  CAGAGCTTGATATGAAATTAACCAAGAAAGAAATTAAGTGGCGTTTATTAATTTTCAA 360
Qy      179  GluGlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeu-----LysGlu 196
Db      361  GAGTTGGTGTAATGCAACTTGTATTGAAACTAGTGTCGCTGATCACTGATTAATAAAT 420
Qy      197  ArgLeuProLysSerPheLysGln----- 204
Db      421  AGTGTCAAGACTTCAGAAAAGCAATTAATTGACAGTAAAGAAATTTTGAAGTGTCT 480
Qy      205  -----AspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp--- 220
Db      481  GAACACCAATATGCGTTCTCACTGTTGCTTGGCTGCTGCTGATTAATTAAGTCCGGAACAG 540
Qy      221  -----AlaMetAlaLeuAspProProAspPheIleThr 231
Db      541  TGGCAGCAAGTCACTACGAAATTCGCCCATATCCCTTACAA-----ACT 585
Qy      232  GlyTyrGluPheValGluGluAlaLeuLysLeuGlnGluGluGlyAlaSerSerLeu 251

```



```

Db      586  GGTCAAGAG-----CTGCTACTAGCTGAAGCT-----TTG 615
Qy      252  AlaProAspLeuArgAlaGlnIleAspGluThrLeuGlnGluIleThrProArgTyrVal 271
Db      616  TTCTCCAGTTCACAGCGCAAAATTCAGGCTGATCTTAACAAAATTCGCGCAATTCGAAT 675
Qy      272  LeuGlnLeuLeuGlnIleuProLeuGlnIleAspAspTyrAlaAlaLysArgLeuAsnGlyLeu 291
Db      676  TTGGAGTGTCTCCGATTACTCT-----CAGAAAAGACTCCGACAGCAAGCAAGCTTGA 729
Qy      292  SerGlyValArgAsnIleLeuTyrPseValGlyGly---GlyGlyAlaSerAlaLeuVal 310
Db      730  GAATTAATTCGCAAAATCTCTAGAAAGATCGGCGGATGATGCGACAGAACAATGATGAA 789
Qy      311  GlyGlyLeuThrArgGlu-----LysPheMetAsnGluAlaPheLeuAspMetThr 327
Db      790  TCGGGTTTAAACATGATGACTTTCGCAATTTATCAGACGATTAACGCAACCACTTACAA 849
Qy      328  AlaAlaGlnGlnValaAspLeuPheValAlaThrProSerAsnIleProAlaGluSerPhe 347
Db      850  GTTCGACAGACGCAAGCTTATTGAAAGCTCAAGCAACGCT-----TCTTCGCTGTT 903
Qy      348  GluValTyrGluValAlaLeuValAlaGlnAlaPheIleGlyLysLysProHis 367
Db      904  GCCACTTAATTAGCTGTTATGCTTATGCTGATAGCGGAGATTTGCTCAACGCAACCTGCT 963
Qy      368  LeuLeuGlnAspAlaAspLysGlnPheGlnGlnIleGlnGlnAlaLysValMetAlaMet 387
Db      964  TTTAATTTGTCAGAGAGACAAATGCTCGTGGCTTGGGCAAGCGCCAA-----1011
Qy      388  GluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGlu 407
Db      1012 -----GATGTAACATTGAA 1026
Qy      408  ArgGlyLeuCysAlaLeuLeuIleGlyLysValaAspGluCysArgMetTrpLeuGlyLeu 427
Db      1027  CAGTCGTATGCTGCTTACTTTGGGCAAACTGAAAGACGACTCGTGTATTGAACTT 1086
Qy      428  AspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPheValLeuGluAsnSer 447
Db      1087  AGTCAGAG-----TAGAA-----GCTTAACTTTATTCGGGAAAAAATCT 1128
Qy      448  AsnArgAspAspAsnAspAspLeuProGlyLeuCysLysLeuLeuGluThrTrpLeuAla 467
Db      1129 -----CAGACTCTCCAGATTGTTACCGGCTGTGTATTATGCAAGAACGCTGCTCA 1185
Qy      468  GlyValValaPheProArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr 487
Db      1186  CAGAACTCTTTCCCAATTTTCGAGATTAGCAAAACGACGAACTTCTTAAAGATTAC 1245
Qy      488  TyrAspAspProMetValLeuSerTyrLeuGlnArgValGluValValGlnGlySerPro 507
Db      1246  TTGCTCAACCAACAGGTGCAAGCTTATTTAGAA-----1278
Qy      508  LeuAlaIleAlaIleThrMetAlaArgIleGlyValaGlnHisValLysAlaSerAlaMet 527
Db      1279 -----GCACTGCCCACTGATGCCCAACAACACTATGATGAGGCTGTAATTAACCCCA 1332
Qy      528  GlnAlaLeuGlnLysValaPheProSerArgTyrThrAspArgAsnSerAlaGluProLys 547
Db      1333  TATTTTCCCAAGCGCAAGCAAGAAATATCATTTTCAATTAACATTAACATTAACATTA 1392
Qy      548 -----AspValGlnGluThrValaPheSerVal 556
Db      1393  GCGTCATTATATCACAGACAGATACCTAACCCAGATTGCGCAAGAAACA-----1440
Qy      557  AspProValGlyAsnAsnValaGlyArgAspGlyGluProGlyValaPheIleAlaGluAla 576
Db      1440 -----1440
Qy      577  ValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGlu 596

```

```

Db      1441 -----CAAACAAA-----GAAACCTTGAAATATCAAACTTTCACACCACTATGTG 1488
Qy      597  SerSerValaAspGluThrThrValaGluMetSerValaAlaAspMetLeuLys-----613
Db      1489  AGTTCAATCGGAAGTATTAATCAAGAGTTCCGTGCTCGAAAGATGACAGAGGATCT 1548
Qy      614 -----GluAlaSerValaLysIleLeuAlaAlaGly-----623
Db      1549  AATCACTATTGAACGGTTCAGCTAAGAGTGCATCTGTGTCATAACAAACCGTAGG 1608
Qy      623 -----623
Db      1609  CGGAGAAAACCTATCTCACTGCTAGCCGAGACGATATACAGATTAATGCTCTCATTTCT 1668
Qy      623 -----623
Db      1669  CGTGTCCCGAGAGCGGGAACCTTTTGGCAACACATAGAAAGTAAACACGCGTGTGA 1728
Qy      624 -----ValAlaIleGlyLeuIleSerLeu-----631
Db      1729  TCGAGAGTGTATTCTTCTTGTGACATATTAAGTTTGTGATATTAAGCCACAACT 1788
Qy      632 -----PheSerGlnLysTyrPheLeuLysSerSerSerPheGlnArgLysAspMet 649
Db      1789  TTGGATGTATAAAATCTGTTTCTTCTCAACCTTTCGCTGATCTACAGTGTGTT 1848
Qy      650  ValSerSerMetGluSerAspValaAlaThrIleGlySerValaArgAlaAspAspSerGlu 669
Db      1849  GTACAAATTAACCAACACACGCTTACCATTCGCGATCCAAATTAAGAAACCAATCGAA 1908
Qy      670  AlaLeuProArgMetAspAlaArgThrAlaGluAsnIleValSerLysTrpGlnLysIle 689
Db      1909  GAAGCCCTTTAACAATGACAGAG---GCAGAAAGATTAATTCACACTTGTTATCTACC 1965
Qy      690  LysSerLeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValaLeuAspGly 709
Db      1966  AAAGCCGACGCTTTAGGCGCCCAATCAAGATTAATTAATTAAGCAAAATTTAACTGCT 2025
Qy      710  ArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyr 729
Db      2026  TCAGCTTATCTCAATGCGCA---CTGATTCCTCAACAGATTAAGTTAGCAATTCGCTAC 2082
Qy      730 -----AspTyrThrLeuLeuLysLeuSerValaAspSerValaThrValaSerAlaAsp 746
Db      2083  CGCAAGTTCACCATAGTTTGAAGATAGAAATCTGTTGAGAAATTTGTTATTTGACAT 2142
Qy      747  GlyThrArgAlaLeuValaGlnAlaThrLeuGlnGluSerAlaCysLeuSerAspLeuVal 766
Db      2143 -----CGTCCCGACATGAGAGCTACGCTCAAGAAAGTACGCACTTATTAAGAAATAT 2196
Qy      767  HisProGluAsnAsnAlaThrAspValaArgThrTyrThrThrArgTyrGluValaPheTrp 786
Db      2197  CAGTTTAAAACTCTTCAACGAT-----AAATTAAGACTGGTATGACTTGAATTGCA 2250
Qy      787  SerLysSerGlyTrpLysIleThrGlnGlySerValaLeu 799
Db      2251  GAACGAGTAATGCGCTATTCAAGATCACTGTTGTA 2289

```

Search completed: February 21, 2006, 13:05:15
 Job time : 1228 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2006, 13:17:07 ; Search time 1565 Seconds
(without alignments)
4232.442 Million cell updates/sec

Title: US-10-600-070B-2

Perfect score: 4063
Sequence: 1 MEALSHVIGLSPFQCRRLP.....YEVFWKSGMKITGSVLA 801

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Qc/abes/ABSSWB.spool/US10600070/rnatc.21022006.112404.8748/app.query.fasta_1
-DB=Published.Applications.NA.Main -OPMT=fastcap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1
-MATRIX=bloms62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs805h
-USER=US10600070.CCEN.1.1.1026.@runatc.21022006.112404.8748 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_NA_Main.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4063	100.0	2406	7	US-10-600-070-1
2	4051	99.7	2406	7	US-10-600-070-128
3	4051	99.7	2637	7	US-10-600-070-130
4	4051	99.7	2679	8	US-10-739-930-127
5	4049	99.7	2406	7	US-10-600-070-9
6	3837.5	94.4	3667	7	US-10-600-070-3
7	3823.5	94.1	3667	7	US-10-600-070-10

8	1775.5	43.7	2283	7	US-10-600-070-126	Sequence 126, App
9	1134	27.9	1146	7	US-10-424-599-129007	Sequence 129007, A
10	973	23.9	1411	8	US-10-425-115-81853	Sequence 81853, A
11	960	23.6	1703	7	US-10-437-963-69933	Sequence 69933, A
12	856	21.1	1536	8	US-10-425-115-57452	Sequence 57452, A
13	784	19.3	561	7	US-10-600-070-132	Sequence 132, A
14	733.5	18.1	631	7	US-10-600-070-184	Sequence 184, App
15	718	17.7	660	7	US-10-600-070-135	Sequence 135, App
16	546	13.4	622	7	US-10-600-070-146	Sequence 146, App
17	541	13.3	537	7	US-10-600-070-143	Sequence 143, App
18	529	13.0	607	7	US-10-021-323-3562	Sequence 3562, App
19	524.5	12.9	2307	7	US-10-600-070-166	Sequence 166, App
20	503	12.4	608	7	US-10-600-070-137	Sequence 137, App
21	501	12.3	552	7	US-10-021-323-3636	Sequence 3636, App
22	499	12.3	1039	7	US-10-424-599-35059	Sequence 35059, A
23	498.5	12.3	2370	7	US-10-600-070-193	Sequence 193, App
24	497	12.2	2400	7	US-10-600-070-163	Sequence 163, App
25	482	11.9	2130	7	US-10-437-963-69932	Sequence 69932, A
26	480	11.8	491	7	US-10-600-070-174	Sequence 174, App
27	474.5	11.7	3554	7	US-10-437-963-53316	Sequence 53316, A
28	472	11.6	446	7	US-10-600-070-151	Sequence 151, App
29	463	11.4	545	7	US-10-600-070-175	Sequence 175, App
30	453.5	11.2	647	7	US-10-600-070-185	Sequence 185, App
31	448	11.0	871	7	US-10-600-070-153	Sequence 153, App
32	446	11.0	2469	7	US-10-600-070-4	Sequence 4, App1
33	446	11.0	2469	7	US-10-600-070-161	Sequence 161, App
34	443	10.9	652	7	US-10-600-070-186	Sequence 186, App
35	435.5	10.7	647	7	US-10-767-701-4069	Sequence 4069, App
36	431	10.6	2857	7	US-10-600-070-172	Sequence 172, App
37	429	10.6	2145	7	US-10-600-070-168	Sequence 168, App
38	425	10.5	2022	7	US-10-600-070-191	Sequence 191, App
39	413	10.2	2511	7	US-10-600-070-189	Sequence 189, App
40	409.5	10.1	1032	7	US-10-767-701-9474	Sequence 9474, App
41	403	9.9	2151	7	US-10-600-070-150	Sequence 150, App
42	391.5	9.6	479	7	US-10-600-070-157	Sequence 157, App
43	390	9.6	460	7	US-10-600-070-187	Sequence 187, App
44	387.5	9.5	439	3	US-09-732-627A-2154	Sequence 2154, App
45	381.5	9.4	420	7	US-10-600-070-176	Sequence 176, App

ALIGNMENTS

RESULT 1
US-10-600-070-1
Sequence 1, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Oseeryoung, Katherine W.
APPLICANT: Vitsha, Stanislav
APPLICANT: Kokeharova, Olga A.
TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 2406
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Alignment Scores:

Pred. No.: 0
Score: 4063.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 7
Length: 2406
Matches: 801
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-600-070B-2 (1-801) x US-10-600-070-1 (1-2406)

QY 1 MercGlualLeuSerHisValGlyLeuSerProPheGlnLeuCybArgLeuPro 20
Db 1 ATGGAAGCTCTGAGTCACTGGCATGTGCTCTCCCATTTCAATTATGCGATTTCACA 60
QY 21 ProAlaThrThrIleuSerLeuArgSerHisAsnThrSerThrIleCybSerAlaSer 40
Db 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAACCTCTCAACTATCTGCTCGGCAGC 120
QY 41 LysThrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
Db 121 AAATGGGCGCAGCGCTCTCTCGACTCAATTTCACCTCCGATTCCTCTCTCTCTCC 180
QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
Db 181 TTGGCCACCGCCACACACCGCACCTCTGCTCTCTGCGACCATATATGATCGTCC 240
QY 81 GluArgHisValProIleProIleAspPheTyrglnValLeuGlyAlaGlnThrHisPhe 100
Db 241 GAACGCCACGTCCCATCCCATTTGATTTCTACCAAGGTATTAGSAGCTCAACCATTTCC 300
QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProProGlnPheGly 120
Db 301 TTAAACGATGAAATCAGAAAGACATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGCT 360
QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCybGluThrLeu 140
Db 361 TTCAGCCACACGCTTTAAATCAGCGGAGACAGATCTTCAGAGCTGCTTGCGAAACTCTG 420
QY 141 SerAsnProArgSerArgArgGluTyrgAsnGluGlyLeuLeuAspAspGluAlaThr 160
Db 421 TCTAATCTCGCTCTAAGAGAGTACATGAAAGCTCTTCTTGATGATGAAGAGCTACA 480
QY 161 ValIleThrAspValProThrAspIleValProGlyAlaLeuCybValLeuGlnGly 180
Db 481 GTCATCACTGATGTTCTTGAGTAAAGTTCCTGGGCTCTCTGTGTATGTCAGAAAGGT 540
QY 181 GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys 200
Db 541 GGTAGAGCTAGATAGTTCTTGCGGGTGTAGGCTCTGCTTAAAGAGAGGTTCCTTAG 600
QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
Db 601 TCGTTTAAAGCAAGATGCTGTTTATGCTTATGCGCTTCTCGATTCTCGAATGCTGAGGAT 660
QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrgIlePheValGluGluAlaLeu 240
Db 661 GCATGCGCATTTGATTCACCTGATTTATCTGCTTATGATGATGTTGTGAGGAAGCTTTG 720
QY 241 LysLeuLeuGlnGluGluGlyValaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
Db 721 AAGCTTTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 261 GluThrLeuGluGluIleThrProArgTyrgValLeuGlnLeuLeuGlyLeuProLeuGly 280
Db 781 GAGCTTTTGAAGAGATCACTCCGCTTATCTTGAAGCTACTTGGTTTACCGCTTGCT 840
QY 281 AspAspTyrgAlaAlaIleValArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
Db 841 GATGATTAACGCTCGGAAAGACTAAATGTTTAAAGCGGTGCGGAAATATTTGTGCTCT 900
QY 301 ValGlyGlyGlyValaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
Db 901 GTTGAGAGAGGTGAGCATGCTCTGTTGGGGGTGTGACCGGTGAGAAATTTATGAT 960
QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValaAspLeuPheValAlaThrProSer 340
Db 961 GAGCGCTTTTACAAATGACACGCTGCTGACAGGTGATCTTTTGTGATGATACCCCAAGC 1020
QY 341 AsnIleProAlaGluSerPheGluValTyrgIleValaAlaLeuAlaLeuValAlaGlnAla 360
Db 1021 AATATTCACAGAGAGTCAATTGAAATTACAAAGTTGACATTGCTCTGTGCTCAAGCT 1080

QY 361 PheIleGlyLysLysAspHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGln 380
Db 1081 TTATATGTAAAGAGCACACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrgAspThrArgAsnAsnTrp 400
Db 1141 CAGGCTAAGGTAAATGGCTATGAGATTCCTGCGATTTGTATGATACAGGAATTAATGG 1200
QY 401 GluIleAspPheGlyLeuGluThrArgGlyLeuCybAlaLeuLeuIleGlyLysValaAspGlu 420
Db 1201 GAGATGACTTCGCTCTAAGAAAGGAGACTCTGTCATCTGCTTATAGCAAAAGTTGATGAA 1260
QY 421 CybArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrgAsnProAlaIleVal 440
Db 1261 TGCCTATGTGTGGCTTAGACATGAGATTCACAATATATGAAATCAGACTATATGTG 1320
QY 441 GluPheValLeuGluLysSerAsnAlaGAspAspAsnAspAspLeuProGlyLeuCybLys 460
Db 1321 GAGTTTGTATGAGAAATTCMAATCGATGACAAATGATGATCTCCCTGACATATGCMAA 1380
QY 461 LeuLeuGluThrTrpLeuAlaGlyValaPheProArgPheArgAspThrLysAspLys 480
Db 1381 TTGTTGAAACCTGCTGTTGGCAGGGGTGCTTCTTCTATGCTTACAGACACCAAGATTA 1440
QY 481 LysPheLysLeuGlyAspTyrgTyrgAspAspPrometValLeuSerTyrgLeuGluArgVal 500
Db 1441 AAATTTAACTCGGGGACTATGATGATCTATGATGTTTGAATTAATTGGAAGAGTGT 1500
QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyValaGlu 520
Db 1501 GAGTGTGTTCAAGGTTCTCTTTAGCTGCTGCAACTATGCGAAGATTTGAGCCGAG 1560
QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValaPheProSerArgTyrgThrAsp 540
Db 1561 CATGTAAACCTGCTGCTATGACGACCTGCAGAAATTTTCTTCCGCTATACAGAT 1620
QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValaAspProValGly 560
Db 1621 AGAACTCGGCTGAACCAAGATGTGCAAGACAGATGTTATGATGATCTCTGTTGCT 1680
QY 561 AsnAsnValGlyArgAspGlyGluProGlyValaPheIleAlaGluAlaValaArgProSer 580
Db 1681 AACAAATGAGCCGTATGAGTGTGAGCTGTGCTTTATGCAAGAGCTGTAAGACCTCT 1740
QY 581 GluAsnPheGluThrAspAspTyrgAlaIleArgAlaGlyValaSerGluSerSerValaAsp 600
Db 1741 GAAATCTTGAACCTAATGATTTATGCAATTCGAGCTCGGCTCAAGAGATACGTTGAT 1800
QY 601 GluThrThrValaGluMetSerValaAspMetLeuLysGluLysValaSerValaLysIleLeu 620
Db 1801 GAAACTACTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
QY 621 AlaAlaGlyValaAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrgPheLeuLysSer 640
Db 1861 GCTGCTGCTGTGCAATYGACATGATTTCACTGTTCAGCCAGAAATTTTCTTAAAGC 1920
QY 641 SerSerSerPheGlnArgLysAspMetValaSerSerMetGluSerAspValaAlaThrIle 660
Db 1921 AGCTCATCTTTTCAACCAAGATATGATTTCTTCTATGAAATCTGATGCTCCTACATA 1980
QY 661 GlySerValaArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
Db 1981 GGGTCACTGAGAGTGAAGATTGAGAGCACTTCCCAATGATGATGCTAGAGCTGACAG 2040
QY 681 AsnIleValaSerLysTyrgIleLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
Db 2041 AATATGATATCAAGTGTGAGAAATTAATGATCTCTGCTTTTGGGCTGTATCACCGCAT 2100
QY 701 GluMetLeuProGluValaLeuAspGlyArgMetLeuLysIleTrpThrAspArgAlaAla 720
Db 2101 GAAATGTTACAGAGGTTTGTGATGGCGAATGCTGAAGATTTGACATGACAGAGCACT 2160
QY 721 GluThrAlaGlnLeuGlyLeuValTyrgAspTyrgThrLeuLeuLysSerValaAspSer 740

```

Db      2161 GAAACTCGCAGCTGGTGGTTATGATATACAGTGTGAACATCTGTTGACAGT 2220
Qy      741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 760
Db      2221 GTGACATCTTCAGCAGATGGAACCGTGCTCTGTGGAGCACTCTGGAGAGCTTGCT 2280
Qy      761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
Db      2281 TGCTATCTGATTTGGTTGTTATCCAGAAACATCTCTATGATGTCAAGACTTACACA 2340
Qy      781 ArgTyrGluValPheTyrSerLysSerGlyTyrLysIleThrGluGlySerValLeuAla 800
Db      2341 AGATACGAAGTTTCTGCTCCAGTCAAGTCAAGGTGGAATAATCACTGAAGCTCTGTTCTTCA 2400
Qy      801 Ser 801
Db      2401 TCA 2403

```

RESULT 2

```

US-10-600-070-128
; Sequence 128, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitcha, Stanislav
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Placetic Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

```

```

Alignment Scores:
Pred. No.: 0          Length: 2406
Score: 4051.00       Matches: 799
Percent Similarity: 99.8%   Conservative: 0
Best Local Similarity: 99.8%   Mismatches: 2
Query Match: 99.7%         Indels: 0
DB: 7                  Gaps: 0

```

US-10-600-070b-2 (1-801) x US-10-600-070-128 (1-2406)

```

Qy      1 MercGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGluLeuCyAspArgPro 20
Db      1 ATGGAAGCTCTGAGTACGTGCGCATGTGCTCTCCCATTCACATTAATCCGATTACCA 60
Qy      21 ProAlaThrThrLysLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
Db      61 CCGGCACGACAAAGCTCCGACGAGCCACAAACCTCTCAACTACTCTCTCCGCGCAGC 120
Qy      41 LysTyrAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
Db      121 AAATGGGCGGACCGTCTCTCTCGACTTCAATTTCACCTCCGATTCTCTCTCTCC 180
Qy      61 PheAlaThrAlaThrThrThrAlaThrLysValSerLeuProProSerIleAspArgPro 80
Db      181 TTGCGCACCGCCACACACCGCCACTCTGCTCTCCGACCATCTATTGATCCGCC 240
Qy      81 GluArgHisValProIleProIleAspPheTyrGluValLeuGlyIleGlnThrHisPhe 100
Db      241 GAAGCGCAGCTCCCATCCCATGATTCTTACCAAGTATTAGAGCTCAAAACATTTC 300
Qy      101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProGlnPheGly 120

```

```

Db      301 TTAACCGATGAAATCAGAGAGCATTCGAGCTTAGGGTTTCGAACCGCCGCAATTCGT 360
Qy      121 PheSerAspAspAlaLeuIleSerArgGlnIleLeuGlnAlaIleCysGluThrLeu 140
Db      361 TTCAGCAGCAGCGTTTAAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTTG 420
Qy      141 SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
Db      421 TCTATCTCTCGCTTAAGAGAGATCAAGAAAGCTTCTTGAATGATGAAGAGCTACA 480
Qy      161 ValIleThrAspValProTyrAspLysValProGlyAlaLeuCysValLeuGlnGly 180
Db      481 GTCATCACTGATGTTCTTGAGATTAAGGTTCTGTGCTCTCTGTATTCAGAAAGGT 540
Qy      181 GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys 600
Db      541 GGTGAGCTAGATGATGTTCTTGGGTGGTGGAGCTTGTCTTAAGAGAGGTTGCTTACG 660
Qy      201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
Db      601 TCGTTTAAAGCAAGATGCTGTTTATGATGCGCTTCTCGATGCTCGAGGGAT 660
Qy      221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
Db      661 GCTATGCAATTTGATCCACTGATTTTAACTGTTATGATGATTTGTTGAGAGAGCTTTG 720
Qy      241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuAspAlaGlnIleAsp 260
Db      721 AACCTTTTACAGAGAGAGAGAGAGAGAGCTTGCACCGGATTTACGTGCAAAATGAT 780
Qy      261 GluThrLeuGluGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGly 280
Db      781 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGAGACTTACTTGCTTACCCCTTGCT 840
Qy      281 AspAspTyrAlaIleLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSer 300
Db      841 GATGATTACCTCGCAAAAGACTAAATGTTTAAAGCGGTGCGGAATTTTGTGCTCT 900
Qy      301 ValGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
Db      901 GTTGAGAGAGTGTGAGATCAGCTCTGTTGGGGTGTGACCGGTGAGAGATTTATTAAT 960
Qy      321 GluAlaPheLeuArgMetThrAlaIleGluGlnValAspLeuPheValAlaThrProSer 340
Db      961 GAGCGCTTTTACGATGACAGCTGCTGAGCAGAGCTGATTTTGTAGCTTACCCCAAGC 1020
Qy      341 AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAla 360
Db      1021 AATATTCCAGCAGAGTCATTGTGAAGTTTACGAAGTTGCACTTCTGTGGCTCAAGCT 1080
Qy      361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
Db      1081 TTTATTGTGAAGAGCACCTTTTACAGAGTCTGATTAAGCAATTCACGAACTTCAG 1140
Qy      381 GluAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTyr 400
Db      1141 CAGGCTTAAGTAAATGGCTATGAGATTCCTGCGATGTTGATATACACGAAATTAATGG 1200
Qy      401 GluIleAspPheGlyLeuGluArgGlyLeuCyAspAlaLeuLeuIleGlyLysValAspGlu 420
Db      1201 GAGATTAAGCTTGGCTTAAGAAAGGGAGCTGTGCACTGTTATAGGCAAAAGTTGATGA 1260
Qy      421 CysArgMetTyrLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
Db      1261 TGCCTGATGTGTGGCTTGAACAGTGAAGATTCACAATATGGAATTCACGCTAATTTG 1320
Qy      441 GluPheValLeuGluLysSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
Db      1321 GAGTTTCTTTTGAAGATTCAAATCGTATGACAAATGATATCTCCCTGACATATGCAAA 1380
Qy      461 LeuLeuGluThrTyrLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
Db      1381 TTGTTGAAACCTGTGGCAGGGTGTGCTTTCTTCAAGTTTCAAGACACCAAAAGTAA 1440

```

```
QY 481 LysPheIyaleuGlyAspTyrTrpAspPProMetValIleuSerTyrIleuGluArgVal 500
DB 1441 AAATTAACTCGGGGCTACTATGATGATCTTATGATTGATTAAGGAGTG 1500
QY 501 GluValValGlnGlySerProIleuAlaAlaAlaIleuAlaArgIleGlyAlaGlu 520
DB 1501 GAGGTAGTTCAGGGTCTCTTATGCTGCTGCTGCAAGCTTGGCAAGATTGGACCGAG 1560
QY 521 HisValIleValIleSerAlaMetGlnAlaIleuGlnIleValIlePheProSerArgTyrThrAsp 540
DB 1561 CATGTGAAGCTAGTGTCTATGACAGGCACTGCAAGAACTTTCTTCCCTATACAGAT 1620
QY 541 ArgAsnSerAlaGluProIleValAspValGlnGluThrValIlePheSerValAspProValGly 560
DB 1621 AGAACTCGGCTGAAACCAAGAGATGCAAGACAGATGTTTATGTTAGTATCCGTGGT 1680
QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
DB 1681 AACAAATGAGGCGGTATGATGAGCTGTGTCTTATTCAGAAAGCTGTAAAGCCCTCT 1740
QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
DB 1741 GAAACTTTGAACTATATGATTATGCAATTGCAAGCTGGGTCTCAGAGATAGCGTTGAT 1800
QY 601 GluThrThrValGluMetSerValAlaAspMetIleuGluIleSerValIleValIleu 620
DB 1801 GAAACTACTGTGTAAGATGTCGTTGATATGTTAAAGAGCAAGTGTGAAGATCTTA 1860
QY 621 AlaAlaGlyValAlaIleGlyIleuIleSerIlePheSerGlnIleValIlePheIleuIleSer 640
DB 1861 GCTGCTGCTGTGGCAATTGACATGATTTCACTGTCCAGCCAGAAATTTTCTTAAAGC 1920
QY 641 SerSerSerPheGlnArgIleValAspMetValSerSerMetGluSerAspValAlaThrIle 660
DB 1921 AGCTCATCTTTTCAACCAAGATATGATGTTCTTCTATGCAATGATGTCGCTACCATTA 1980
QY 661 GlySerValArgAlaAspAspSerGluAlaIleuProArgMetAspAlaArgThrAlaGlu 680
DB 1981 GGGTCATCTCAGAGCTGACATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGAG 2040
QY 681 AsnIleValSerIleValIleArgIleValIleSerSerIleuAlaPheGlyProAspHisArgIle 700
DB 2041 AATATGATATCCAAAGTGCAGAAATTAATGATCTGTGCTTTTGGCCGTGATCAACGCATA 2100
QY 701 GluMetLeuProGluValIleuAspGlyArgMetIleuValIleTyrThrAspArgAlaAla 720
DB 2101 GAATGTTACAGAGGTTTGGATGGGCGAATGCTGAAGATTGGAAGCTGCAGAGCAGCT 2160
QY 721 GluThrAlaGlnIleuGlyIleuValTyrAspTyrThrIleuIleuValIleAspSer 740
DB 2161 GAAACTGCGCAGCTTGGGTGGTTATGATTATACCTGTTGAACATATCTGTGACAGT 2220
QY 741 ValThrValSerAlaAspGlyThrArgAlaIleuValGluAlaThrIleuGluIleuSerAla 760
DB 2221 GTGACATCTTCAGAGATGAAACCCGCTCTGCTGGAAGCACTCTGGAGAGATCTGCT 2280
QY 761 CysIleuSerAspLeuValIleAspProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
DB 2281 TGCTATCTGATTTGGTTCACTCAGAAACAAATGCTACTGATGTCAGAACTTACACACA 2340
QY 781 ArgTyrGluValIlePheTyrPheSerIleSerGlyTyrIleValIleThrGluIleuValIleuAla 800
DB 2341 AGATACCAAGTTTCTGTCTCAGATGCAAGGTGAGAAATCACTGAAGGCTGTGTTCTTGCA 2400
QY 801 Ser 801
DB 2401 TCA 2403
```

RESULT 3
US-10-600-070-130
; Sequence 130, Application US/10600070
; Publication No. US20040139500A1

```
; GENERAL INFORMATION:  
; APPLICANT: Osteoryong, Katherine W.  
; APPLICANT: Vitha, Stanislaw  
; APPLICANT: Kokharova, Olga A.  
; APPLICANT: Gao, Hongbo  
; TITLE OF INVENTION: Platelet Division and Related Genes and Proteins, and Methods of  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 130  
; LENGTH: 2637  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-600-070-130  
  
Alignment Scores:  
Pred. No.: 0 Length: 2637  
Score: 4051.00 Matches: 799  
Percent Similarity: 99.8% Conservative: 0  
Best Local Similarity: 99.8% Mismatches: 2  
Query Match: 99.7% Indels: 0  
DB: 7 Gaps: 0  
  
US-10-600-070B-2 (1-801) x US-10-600-070-130 (1-2637)  
  
QY 1 MetGluAlaIleuSerHisValGlyIleGlyIleuSerProPheGlnIleuCyArgLeuPro 20  
DB 114 ATGAAGCTCTGAGTACGCGGCAATGGTCTCTCCCAATTCAAATTATGCAATTATACCA 173  
QY 21 ProAlaThrThrIleuValArgSerHisAsnThrSerThrIleCysSerAlaSer 40  
DB 174 CCGGCAACGACAAAGCTCCAGACTAGCCACACACACTCTACAACTATCTCTCCGACG 233  
QY 41 LysTrpAlaAspArgLeuIleuSerAspPheAsnPheThrSerAspSerSerSerSer 60  
DB 234 AATAGGCGCAGCCGTCCTCTCTCCGACTTCAATTCACCTCCGATTCCTCTCTCC 293  
QY 61 PheAlaThrAlaThrThrThrAlaThrIleuValIleuProProSerIleAspArgPro 80  
DB 294 TTGCGCCACCGCCACACACACGCGCCTGCTCTCCGACACACTTATGATGATCGTCC 353  
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValIleuGlyAlaGlnThrHisPhe 100  
DB 354 GAAAGCCAGTCCCCCATCCCATTTGATTTCTACAGGATATGAGAGCTCAACACATTC 413  
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleuProProGlnPheGly 120  
DB 414 TTAAACGATGATTCAGAAAGAGCATTCGATGAGGTTTCGAAACCGCCGCAATTCGGT 473  
QY 121 PheSerAspAspAlaIleuIleSerArgArgGlnIleuGlnIleuAlaIleuGlnThrIleu 140  
DB 474 TTCAGGACACACGCTTATACGCGGAGACAAATCTTCAACCTGCTTCCGAAACTCTG 533  
QY 141 SerAspProArgSerArgArgGlyTyrAsnGlyIleuLeuAspAspGluGluAlaThr 160  
DB 534 TCTAATCTCGGTCTAAGAGAGTACATGAAAGTTTCTTATGATGATGAAAGAGTACA 593  
QY 161 ValIleThrAspValProTyrAspValValProGlyAlaIleuCyValIleuGlnGly 180  
DB 594 GTCATCTGATGTTCTTGGATTAAGTTCTGTGCTCTCTGTGATTCGAAAGAGT 653  
QY 181 GlyIleuThrGluIleuValIleuValGlyGluAlaIleuIleuValGluGluProIle 200  
DB 654 GGTGAGACTGAGATGTTCTTCCGGTTGAGAGGCTGTGCTTAAAGAGAGTCCCTAAG 713  
QY 201 SerPheIleGlnAspValIleuValMetAlaIleuAlaPheIleuAspValSerArgAsp 220  
DB 714 TCGTTTAAAGCAAGATGAGTTTATGATGAGCGCTTTCGTTTCTCGATGTCGAGAGAT 773  
QY 221 AlaMetAlaIleuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaIleu 240
```

Db	774	GCTAATGGCAATTCGACTTATTTATTAACCTGGTATGACTTTGTTGAGCAACCTTTG	833
Oy	241	LYSLEULENGINGLUGLUYALASERSEULENALAPROASPLEUARGALAGINILEASP	260
Db	834	AAGCTTTTACAGAGAGAGAGCAAGTACCTTTGACCGGATTTTACGTGCACAAATTGAT	893
Oy	251	GIUHTHLEUGLUGLUILLERHPROARGTYRVALLLEUGLEULENGIYLEUPROLEUGLY	280
Db	894	GAGACTTTGGAGAAGATCATCTCCGGTATATGCTTGGAGCTACCTTGGCTTACCCGTTGGT	953
Oy	281	ASPAPRYTALAALELYARGLEUBNGIYLEUSERGIYVALARGASPIILEUETRPSER	300
Db	954	GATGATTTACGCTGCGCAAAAGACTTAATGGTTTTAAAGCGGTGTCGGAATATTTTGCGCT	1013
Oy	301	VALGIYGIYGIYVALASERLALAUVALGIYGIYLEUTHARGGLULYPHEMETASN	320
Db	1014	GTTGAGAGAGGGTGGAGCATCAGCTCTTGTTGGGGGTTTGACCCGTGAGAGATTTATGAA	1073
Oy	321	GLUALAPHELEUARGMETHTHALAAGUGINVALAPLEUPHEVALALATHPROSER	340
Db	1074	GAGGGGTTTTTACGATATGATCAGCTCTGAGCAGTTGATCTTTTGTAGCTACCCCAAGC	1133
Oy	341	ASNILEPROALAGLUSERPHEGLUVALTYRGIUVALLALEUALALEUVALALGINALA	360
Db	1134	AATATTTCCAGCAGACTCATTTGAAATTTACGAAAGTTGACCTTGCTTGTGTGCTCAAGCT	1193
Oy	361	PHLEILEGIYLYSLYPROHI;SLEULENGIINAPALAPAPLYSGINPHEGINGLINLEUN	380
Db	1194	TTTATTTGGTAAAGAACCCACCTTTTACAGATGCTGATMAAGCATTCACAGCACTTCAAG	1253
Oy	381	GLINALYRVALMETALAMEGLUILLERPROALAMELEUYRASPETHARGASNAINTP	400
Db	1254	CAGGCTTAAGTAAATGCTATGAGAGATTCCTGCGATGTTGATGATACAGGAAATATTTGG	1313
Oy	401	GIULIEAPHPHEGLYLEUGLUARGGLIYLEUCYBALALEULEULLLEGIYLYRVALLASPLU	420
Db	1314	GAGATAGACTTCGGCTCTAGAAAGGGAGACTCTGTGACGCTTAAAGGCAAGTTGATATAA	1373
Oy	421	CYSARGMETTPLEUGIYLEUAPSPERGILUAPSPERGINTYRARGASNPORALALLEVAL	440
Db	1374	TGCCCTATGCTGGTTGGCTCTTAGACAGTGAAGATTACAAATATAGAAATCCAGACTTTGTG	1433
Oy	441	GIUPHEVALLEUGLUASNSEIRASNAARGAPAPASNAAPLEUPROGLIYLEUCYALYS	460
Db	1434	GAGTTTGTTTGGAGAAATTCAAATCGTGATGACATGATGATCTCCCTGAGCTATGCAAA	1493
Oy	461	LEULENGIUTHRTTPLEUALAGIYVALVALPHEPROARGPHEARGASPTHLYASAPLYS	480
Db	1494	TTGTTGGAAACCTGGTTGGCAGGGGTTGCTCTTCTAGATTACAGACCAACCAAGATATTA	1553
Oy	481	LYSPHELYSLEUGIYASPTTYTYRASPAPSPROMETVALLEUSERTYRLEUGLUARGVAL	500
Db	1554	AAATTTTAAACTCGGGGACTTACTATGATATCCATATGTTTGAAGTACTTGGAAGAGATG	1613
Oy	501	GIUVALVALGINGIYSEPROLEUALAALAAALATHMETALAAAGIILEGIYVALAGLU	520
Db	1614	GAGGTAGTTCAAGGGTTCCTTTAGCTCTGTGTGACGTATGCGAAGATTTGGAGCCGAG	1673
Oy	521	HIVALLYVALASERLALAMEGLUALALEUGLINYSVALPHEPROSEARXTYRTHRASP	540
Db	1674	CATGGAAGACTAGTGCATATGACAGCACATGCAAGAAATTTTCCCTCCCGCTAATACAGAT	1733
Oy	541	ARGASNSERLALAGLUPROLYASAPVALGINGIUTHRTVALPHESEVALAASPPOVALAGLY	560
Db	1734	AGAAATCTGGCTGAACCCAGAGATGTGCAAGAGACAGTTTATGTTAGTATGATCTCGTGTGT	1793
Oy	561	ASNANVALGIYARGASPIYGIUPROGLIYVALPHEILEALAGLUALAVALARPROSER	580
Db	1794	AAACATATGAGGCCGTGATGTGTGAGCTGTGTCTTTATTCGCAAGACGTAAAGACCTCTT	1853
Oy	581	GLUENPHEGLIUTHRTASNAPSTYRALALLEARGALAGIYVALSERGLUSERSEVALASP	600

Db	1854	GAAGAACTTGGAAACCTAATGATTATTCAGATTCAGAGCTGGGCTCTCAGAGAGTACGCTTGAT	1913
Qy	601	GIunthrtHrValIGlumeSeValAlaapMetLeuYgJuLaSeValIlySIIeLeu	620
Db	1914	GAACCTACTGTTGAATATGTCGTTCTGTATGTATTAAGAAGGCAAGTGAAGATCTTA	1973
Qy	621	AlaAlaGlyValAlaIIGlyLeuIleSerLeuPheSerGlnIlyrYrHeLeuIlySer	640
Db	1974	GCTGTGGTGGGCAATGAGATTCGATTTTCACGTTCAGGCAGAAAGATTTTCTTAAAGC	2033
Qy	641	SerSerSerPheGlnArgIySaPheMetValSerSerMetGluSerAspValAlaTrIle	660
Db	2034	AGCTATCTTTTCAACGACGAGATATGTTTCTTCTATGGAATCTGATGTCGTACATA	2093
Qy	661	GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu	680
Db	2094	GGGTCAGTACAGCTGACGATTCAGAAAGCATCTCCAGAAATGGAATGCTAGACTGCAGAG	2153
Qy	681	AsnIleValSerIyETrGlnIlySIIeIySerLeuAlaPheGlyProAspHisArgIle	700
Db	2154	AATATAGATATCCAAAGTGGCAGAAAGATTAACTCTCGCTTGGGCTTGATCACCGCATA	2213
Qy	701	GluMetLeuProGluValLeuAspGlyArgMetLeuIlySIIeTrpHrAspArgAlaIa	720
Db	2214	GAATATTTACAGAGAGTTTGGATGGCCAAATGCTGAAGATTTGACATACAGAGAGCT	2273
Qy	721	GIunthralaGlnLeuGlyLeuValYrAspYrYrHrLeuLeuIlySeValAspSer	740
Db	2274	GAACCTGGCGACGCTTGGGTGGTTATGATTATACACTGTTGAACCTATCTGTGCACT	2333
Qy	741	ValThrValSerAlaAspGlyYrThrArgAlaLeuValGluAlaThrLeuGluGluSerAla	760
Db	2334	GTCACAGCTCAGCAGATGGAACCCGCTCTGCTGGTGGAAAGCAACTCGAGGAGTCTGCT	2393
Qy	761	CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrYrYrHr	780
Db	2394	TGCTATCTGATTTGGTTCATCCAGAAACAATGCTACTGATGATCAGAACCTACACACACA	2453
Qy	781	ArgYrGluValPheTrpSerIySeSerGlyYrTrpIySIIeHrGluGlySerValLeuAla	800
Db	2454	AGATACGAAGTTTCTGATCCAAAGTCAGGATGGAAGAAATCAGTGAAGGCTCTGTTCTGCA	2513
Qy	801	Ser 801	
Db	2514	TCA 2516	
RESULT 4			
US-10-739-930-227			
; Sequence 227, Application US/10739930			
; Publication No. US20040216190A1			
; GENERAL INFORMATION:			
; APPLICANT: Kovalic, David K.			
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH			
; FILE REFERENCE: 38-21(53377)B			
; CURRENT APPLICATION NUMBER: US/10/739,930			
; CURRENT FILING DATE: 2003-12-18			
; NUMBER OF SEQ ID NOS: 11088			
; SEQ ID NO 227			
; LENGTH: 2679			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
; FEATURE:			
; OTHER INFORMATION: Clone ID: ABATH-23APR03-CLUSTER13643_1			
US-10-739-930-227			
Alignment Scores:			
Pred. No.: 0			
Score: 4051.00			
Percent Similarity: 99.84			
Best Local Similarity: 99.84			
Query Match: 99.74			
DB: 8			
Gaps: 0			
Length: 2679			
Matches: 799			
Conservative: 0			
Mismatches: 2			
Indels: 0			

US-10-600-070b-2 (1-801) x US-10-739-930-227 (1-2679)

QY 1 MetGluAlaLeuSerHisValIglYlIeGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 Db 115 ATGGAGAGCTCTGAGTACGTCGGCATTTGGTCTCTCCCATTTCCAAATTAATGCCCATTACCA 174
 QY 21 ProAlaThrThrIleuLeuArgArgSerHisAsnThrSerThrThrIleCySerAlaSer 40
 Db 175 CCGGCAACGACAAAGCTCCGACGTAGCCACAAACCTCTTCAACTATCTGCTCCGCCAGC 234
 QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerArgSerSerSerSer 60
 Db 235 AATAGGCGCGACGCTGCTCTCCGACTTCATTTGCACTCCGATTCCTCTCTCTCC 294
 QY 61 PheAlaThrAlaThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 Db 295 TTCGCCACCGCCACCAACCGCCACTCTCTCTCCGCCACCATCTATTTGATCGTCC 354
 QY 81 GluArgHisValProIleProIleAspPheTyArgIleValIleuGlyAlaGlnThrHisPhe 100
 Db 355 GAAGCGCAGTCCCATCCCATTTGATTTCTACCAAGTATTAGAGCTCAAAACAATTC 414
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
 Db 415 TTAAACGATGAGATCAGAAAGCATTCGAAAGCTAGGGTTTGAACCGCGCAATTCGGT 474
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCyArgIleuThrLeu 140
 Db 475 TTCAGGACGACCGCTTTAATCAGCCGACGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 534
 QY 141 SerAsnProArgSerArgArgGlyTyArgAsnIleGlyLeuLeuAspAspGluAlaThr 160
 Db 535 TCTAACTCTCGGCTTGAAGAGATGACAAAGAGCTTTCTTGAATGAAAGACACCA 594
 QY 161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuCyValIleuGlnGly 180
 Db 595 GTCATACGTGATGTTCTTGAGTAAGGTTCTGTGCTCTCTGTGTAATGCAAGAGGT 654
 QY 181 GlyGluThrGluIleValIleuArgValIglYlAlaLeuLeuLysGluValLeuProLys 200
 Db 655 GGTGAGACTGAGATGATGTTCTTGGGTTGAGAGGCTCTGCTTAAAGAGAGTGGCTTAA 714
 QY 201 SerPheLysGlnAspValIleuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 Db 715 TCGTTTAAAGCAAGTGTGTTTAAATGAGCGCTTGTCTCAATGCTCGAGGAT 774
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyArgIlePheValGluAlaLeu 240
 Db 775 GCTATGGCATTTGATCCACTGATTTTAACTGATTATGATGATTGTTGAGGAAGCTTTG 834
 QY 241 LysLeuLeuGlnGluGlyAlaSerSerLeuAlaProAspLeuAlaGlnIleAsp 260
 Db 835 AACCTTTTACAGAGAGAGCAAGAGAGAGAGCTTGCACCGGATTTACGTGACCAATTTGAT 894
 QY 261 GluThrLeuGluGluIleThrProArgTyValIleuGluLeuLeuGlyLeuProLeuGly 280
 Db 895 GAACCTTTGAAAGAGATCACTCCGCTTATGTCTTGAAGCTACTGGCTTACCGCTTGT 954
 QY 281 AspAspTyArgAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 Db 955 GATGATTAAGCTCGAAGAAAGCTTAATGTTTAAAGCGGTGCGGAATATTGTGTGCT 1014
 QY 301 ValIglYlGlyIglYlAlaSerAlaLeuValIglYlLeuThrArgGluLysPheMetAsn 320
 Db 1015 GTTGGAGAGAGTGAAGATCACTCTTGTGGGGTGTGACCCCTGGAAGATTATGAAT 1074
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaIleProSer 340
 Db 1075 GAGCGCTTTTACAAATGACAGCTGCTGAGAGAGTATCTTTTGTAGCTTACCCCAAGC 1134
 QY 341 AsnIleProAlaGluSerPheGluValTyArgIleValAlaLeuAlaLeuValAlaGlnAla 360

Db 1135 AATATTCCAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGCTCAAGCT 1194
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnIleuGln 380
 Db 1195 TTTATTTGGTAAAGCCACACTTTTACAGATGCTGATTAACCATTTCCAGCACTTTCAG 1254
 QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyArgThrArgAsnAsnTrp 400
 Db 1255 CAGGCTAAGATGATGCTATGAGATTCCTGCGATGTTGTATGATACAGGAATATATTGG 1314
 QY 401 GluIleAspPheGlyLeuGluLysArgIleLeuCyValAlaLeuLeuIleGlyLysValAspGlu 420
 Db 1315 GAGATAGACTTGCTGTATGAAGAGGACTCTGCACTGCTTATAGCCAAAGTTATGAA 1374
 QY 421 CyArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyArgAsnProAlaIleVal 440
 Db 1375 TGCCGATATGGTTGGGCTTAAACAGTAGATTCACATATAGGAATCCAGCTATTGG 1434
 QY 441 GluPheValIleuGluAsnSerAsnArgAspAspAsnAspLeuProGlyLeuCyLys 460
 Db 1435 GAGTTTGTGTAAGAAATCAATCGTATGACAAATGATGATCTCCCTGCACTATGCAAA 1494
 QY 461 LeuLeuGluThrTrpLeuAlaGlyValIlePheProArgPheArgAspThrLysAspLys 480
 Db 1495 TTGTTGGAACCTGTTGGCAGGAGTGTCTTCTTCAAGTTCAAGACACCAAGATTA 1554
 QY 481 LysPheLysLeuGlyAspTyTyArgAspAspProMetValLeuSerTyArgLeuArgVal 500
 Db 1555 AATTTTAACTCGGCGGACTATATGATGATCTTATGTTTGAAGTACTTGAAGAGAGT 1614
 QY 501 GluValIleGlnGlySerProLeuAlaAlaAlaIleThrMetAlaArgIleGlyAlaGlu 520
 Db 1615 GAGGTAGTTCAGGGTCTCTCTTACGCTGCTGACGATATGCAAGATTGAGCCGAG 1674
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyTrpAsp 540
 Db 1675 CATGTAAAGCTAGAGCTATAGAGGACGTGCAAGAAAGTTTCTTCCGCTATACAGAT 1734
 QY 541 ArgAsnSerAlaGluProLysAspValGlnIleuThrValPheSerValAspProValGly 560
 Db 1735 AGAACTCGGCTGAACCCAGATGTGCAAGAGACAGTCTTATAGTATATCCTGTTGGT 1794
 QY 561 AsnAsnValIglYArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 Db 1795 AACCATGTAGGCGGTAGTGTGAGCTGTGCTTATATGCAAGACCTTAAGACCTCTGT 1854
 QY 581 GluAsnPheGluThrAsnAspTyArgAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 Db 1855 GAAACCTTGAACCTATATATATGCAATTCAGCTGGGGTCTCAGAGAGTACGTTGAT 1914
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu 620
 Db 1915 GAACCTACTGTTGAATATGTCCTGCTGATATGTTAAAGAGCCAAAGTGTGAAGACTCTTA 1974
 QY 621 AlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyPheLeuLysSer 640
 Db 1975 GCTGCTGAGTGGCAATTTGACTGATTCACCTGTCAGGCAAGATATTTCTTAAAGC 2034
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
 Db 2035 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATAGATCTATGTCCGTACACTA 2094
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 Db 2095 GGGTCACTCAGACTACATTCATTAAGACACTTCCAGAAATGATCTGAGACTGCAAGAG 2154
 QY 681 AsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
 Db 2155 AATATATGATCCAAAGTGGCAGAGATTAAGTCTCTGCTTTTGGCCTGATCAACCCATA 2214
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleThrThrAspArgAlaAla 720
 Db 2215 GAATGTATACAGAGGTTTGTGATGGCGAATGCTGAAGATTTGACTGACAGACAGCT 2274

QY 721 G|uThrA|aG|nLeuG|yLeuVal|TyxAsp|TyThr|Leu|Leu|yLeuSer|Val|AspSer 740
 |||||
 Db 2275 GAA|CTG|CGA|CGCTTGG|TTGG|TTTAT|ATAT|ATAC|CTTTGAA|ACTAT|CTGTTG|ACAGT 2334
 QY 741 Val|Thr|Val|Ser|Ala|Asp|G|yThr|Arg|Ala|Leu|Val|G|u|a|a|Thr|Leu|G|u|G|u|Ser|Ala 760
 2335 GTG|ACAGT|CTCAG|CAGAT|GAA|CCG|TCTCTG|GTGGA|AGCA|CTCTCG|AGAGT|CTGCT 2394
 Db 761 Cys|Leu|Ser|Asp|Leu|Val|His|Pro|G|u|Asn|Ala|Thr|Asp|Val|A|g|Thr|TyThr|Thr 780
 2395 TGT|TAT|CTG|ATTTG|ATTTG|TTCA|TCAG|AAAC|AAT|GCTAT|GAT|G|CAG|AACT|TAC|ACACA 2454
 QY 781 Arg|TyG|u|Val|Phe|Trp|Ser|Leu|Ser|G|yTrp|Lys|Ile|Thr|G|u|G|u|Ser|Val|Leu|Ala 800
 |||||
 Db 2455 AGA|TAC|AAG|ATTTT|CTG|TCTG|ATC|CAAGT|CAAGG|TGA|GAA|ATCA|CTGAA|GGCT|CTGTT|CTT|GCA 2514
 QY 801 Ser 801
 |||||
 Db 2515 TCA 2517
 RESULT 5
 US-10-600-070-9
 ; Sequence 9, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitha, Stanislaw
 ; APPLICANT: Koksharova, Olga A.
 ; TITLE OF INVENTION: Gao, Hongbo
 ; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 2406
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-600-070-9
 Alignment Scores:
 Pred. No.: 0 Length: 2406
 Score: 4049.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 7 Gaps: 0
 US-10-600-070b-2 (1-801) x US-10-600-070-9 (1-2406)

QY 101 Leu|Thr|Asp|G|y|I|e|a|r|G|A|Phe|G|u|a|a|a|r|G|Val|Ser|Lys|Pro|G|In|Phe|G|y 120
 Db 301 TTT|AAC|CAATG|GAATCA|AAG|CAATTTG|AAAG|CTT|AGG|CTTTT|CGAA|CCGCC|CAATTC|GCT 360
 QY 121 Phe|Ser|Asp|Asp|Ala|Leu|I|e|Ser|Arg|G|n|I|e|Leu|G|n|a|a|a|Cys|G|u|Thr|Leu 140
 Db 361 TTC|AGC|A|CAG|C|G|CTT|TAA|TCAG|CCG|GAG|CAG|ATTTCTT|CAAG|CTG|CTCG|GAA|ACT|CTG 420
 QY 141 Ser|Asp|Pro|Arg|Ser|Arg|Val|Tyx|Asp|G|u|y|Leu|Leu|Asp|Asp|G|u|G|u|Ala|Thr 160
 Db 421 TCT|AAT|CTG|G|CTT|TAA|AAG|AGT|TAC|AAG|AGCTT|CTT|CTG|ATG|ATG|AAG|AGCTT|CA 480
 QY 161 Val|I|e|Thr|Asp|Val|Pro|Trp|Asp|Lys|Val|Pro|G|y|Ala|Leu|Cys|Val|Leu|G|n|G|y 180
 Db 481 GTC|ATC|ATG|ATG|TCTT|TGG|AT|AAG|TTC|CTG|GGC|CTCT|CTG|TAT|TTC|AAG|AAG|T 540
 QY 181 G|y|G|u|Thr|G|u|I|e|Val|Leu|Arg|Val|G|y|G|u|a|Leu|Leu|y|G|u|a|Arg|Leu|Pro|Lys 200
 Db 541 GGT|GAG|ACT|GAG|ATG|ATTTCTT|CGGGT|GTG|AGG|CTCTG|CTTAA|GAG|AGTTC|CTAAG 600
 QY 201 Ser|Phe|Lys|G|In|Asp|Val|Val|Leu|Val|Met|Ala|Leu|Ala|Phe|Leu|Asp|Val|Ser|Arg|Asp 220
 Db 601 TCG|TTTAA|GCA|AAG|ATG|GTTTAA|GTAT|G|CCG|TTC|G|TTTCT|CAAT|G|CTTCA|AGG|AT 660
 QY 221 Ala|Met|Ala|Leu|Asp|Pro|Pro|Asp|Phe|I|e|Thr|G|y|Trp|G|u|Phe|Val|G|u|G|u|Ala|Leu 240
 Db 661 GCT|ATG|CAATTTG|ATG|CACTG|ATTTTAT|TAT|CTG|TTATG|ATG|TTTGT|TGA|GAA|GCTTTG 720
 QY 241 Lys|Leu|Leu|G|n|G|u|G|u|G|y|Ala|Ser|Ser|Leu|Ala|Pro|Asp|Leu|Arg|Ala|G|n|I|e|Asp 260
 Db 721 AAG|CTTTTAC|AGG|AGAA|GAG|CAAG|TAC|GCTTGC|ACCG|GATTTAC|CTG|CA|AAATTTG|AT 780
 QY 261 G|u|Thr|Leu|G|u|G|u|I|e|Thr|Pro|Arg|Tyx|Val|Leu|G|u|Leu|Leu|G|y|Leu|Pro|Leu|G| 280
 Db 781 GAG|ACTTTG|AAG|AAG|ATC|CTCG|CGTATG|CTTGA|GCTTAC|TTCG|CTTAC|CGCTTGT 840
 QY 281 Asp|Asp|Tyx|Ala|Ala|Lys|Arg|Leu|Asn|G|y|Leu|Ser|G|y|Val|Arg|Asn|I|e|Leu|Trp|Ser 300
 Db 841 GAT|GATTA|CGCTG|CGAA|AAG|CTTAA|TGT|TTTAA|GCG|TTC|GCG|AAT|ATTTT|GCT|CT 900
 QY 301 Val|G|y|G|y|G|y|Ala|Ser|Ala|Leu|Val|G|y|I|e|Thr|Arg|G|u|y|Sph|Met|Asn 320
 Db 901 GTT|GAG|AGAGT|GAG|ATCAG|CTCTGTTG|GGGGT|TGA|CCG|TGA|AGATTAT|GAT 960
 QY 321 G|u|u|Ala|Phe|Leu|Arg|Met|Thr|Ala|G|u|G|n|Val|Asp|Leu|Phe|Val|Ala|Trp|Ser 340
 Db 961 GAG|CGCTTTTAT|GAA|TGA|CAG|CTG|CGA|CGATG|ATCTTTTGT|TGA|CTAC|CCCA|AGC 1020
 QY 341 Asn|I|e|Pro|Ala|G|u|Ser|Phe|G|u|Val|Tyx|G|u|Val|Ala|Leu|Ala|Leu|Val|Ala|G|n|Ala 360
 Db 1021 AAT|ATTTCC|AAG|ATG|ATTTGA|AGTTTGA|AGTTTGA|AGTTTGA|AGTTTGA|AGTTTGA|AGTT 1080
 QY 361 Phe|I|e|G|y|Lys|Pro|His|Leu|Leu|G|n|Asp|Ala|Asp|Lys|G|In|Phe|G|n|Leu|G|n 380
 Db 1081 TTT|ATTG|TAA|GAG|CAAC|CTTTTAA|CAG|ATG|CTG|ATG|AAG|ATTC|GAG|CAATTC|CAG 1140
 QY 381 G|n|Ala|Lys|Val|Met|Ala|Met|G|u|I|e|Pro|Ala|Met|Leu|Tyx|Asp|Thr|Arg|Asn|Trp 400
 Db 1141 CAG|GCTTAA|GGTAA|TGG|CTATG|AGATTC|CG|CAATG|TTG|TATG|ATAC|ACG|AAATTAATTTG 1200
 QY 401 G|u|I|e|Asp|Phe|G|y|Leu|G|u|Arg|G|y|Leu|Cys|Ala|Leu|Leu|I|e|G|y|Lys|Val|Asp|G|u 420
 Db 1201 GAG|ATG|ATCTT|CGCTTGA|AAGG|GAG|CTTGTG|CACTG|CTTAT|TGA|GCA|AAGTTGATGAA 1260
 QY 421 Cys|Arg|Met|Trp|Leu|G|y|Leu|Asp|Ser|G|u|Asp|Ser|G|n|Tyx|Arg|Asn|Pro|Ala|I|e|Val 440
 Db 1261 TGC|GAT|ATG|G|TGG|GCTTAA|CAG|GAG|ATTC|CAATAT|TGA|ATC|CAATATTTG 1320
 QY 441 G|u|Phe|Val|Leu|G|u|Asn|Ser|Asn|Ala|G|Asp|Asn|Asp|Leu|Pro|G|y|Leu|C|y|Lys 460
 Db 1321 GAG|TTTGT|TTTGA|GATTC|AAATCTG|ATG|TCA|AATG|ATG|ATCTCC|TGA|CATATG|CAAA 1380
 QY 461 Leu|Leu|G|u|Thr|Trp|Leu|Ala|G|y|Val|Phe|Pro|Arg|Phe|Arg|Asp|Thr|Lys|Asp|Lys 480

```

Db      1381 TTGTTGAAACCTGGTTGGCAGGGGTTGTCTTTCCTAGCTTCAGAGCACCAAGATAA 1440
Qy      481 LyePheIyLeuGIyApyTyTyraAPAPPrometValLeuSerTyLeuGIyAglVal 500
Db      1441 AAATTTAACTCGGGGACTATGATGATCTTATGATCTTTTGAAGTCTTGGAAAGAGTG 1500
Qy      501 GluValValGIingIySerProLeuAlaAlaAlaIathrMetAlaArgIleGIyAglVal 520
Db      1501 GAGGTAGTCAAGGTTCTCTTTCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy      521 HisValIyValSerAlaMetGIyAlaLeuGIyValIyPheProSerArgTyTyraAP 540
Db      1561 CATGTGAAGCTAGTGTCTATGACAGGCACTGACAGAAATTTTCTTCCCTATACAGAT 1620
Qy      541 ArgAPSerAlaGIyProIyPheAPValGIyGIyThraValIyPheSerValAPProValGIy 560
Db      1621 AGAAACTCGGCTGAAACCAAGATGTCAGAGACAGTGTTCAGTGTTCCTGTTGCT 1680
Qy      561 AAPASnValGIyArgAPGIyGIyProGIyValPheIleAGIyAlaValArgProSer 580
Db      1681 AACATATAGCGCTGATGGAGACCTGGTGTCTTATGCAAGACTGTAGACCTCT 1740
Qy      581 GluAPNheGIyThraAPAPTyTyraIleArgIleGIyValSerGIySerSerValAP 600
Db      1741 GAAAACTTGAACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy      601 GluThraValGIyMetSerValAlaAPMetLeuIyGIyAlaSerValIyIleu 620
Db      1801 GAAACTATCTGTGAAAGTCCGTTCGTATGTATTAAGAGGCAAGTGTGAAGATCTTA 1860
Qy      621 AlaIleGIyValAlaIleGIyLeuIleSerIyPheSerGIyIyTyPheLeuIySer 640
Db      1861 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy      641 SerSerSerPheGIyArgIyAPAPMetValSerSerMetGIySerAPValAlaIthrIle 660
Db      1921 AGCTCATCTTTCACGCAAGATATGTTCTTCTATGATGATGATGATGATGATGATGAT 1980
Qy      661 GlySerValArgAlaAPAPSerGIyAlaLeuProArgMetAPAlaIthrIle 680
Db      1981 GGGTCAGTCAAGCTGACGATTCAGAAAGCACTTCCGAAATGATGATGATGATGATGAT 2040
Qy      681 APNleValSerIyTyPheIyIleValSerIyLeuIleAPheGIyProAPPhIArgIle 700
Db      2041 AAATATATTCAGATGCGAGAGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy      701 GluMetLeuProGIyValIleuAPGIyArgMetLeuIyIleTyPheIyAPAPArgAla 720
Db      2101 GAAATGTTACAGAGCTTTGGATGGGCGAATGCTGAAGATTTGGACTGACAGAGCAGCT 2160
Qy      721 GluThraIleGIyLeuValTyraAPTyTyraIleLeuIyLeuIyLeuIyLeuIyLeu 740
Db      2161 GAAACTCGGAGCTTTGGTGTATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2220
Qy      741 ValIThrValSerAlaAPGIyThraValIleuValGIyAlaThrLeuGIyLeuIyLeuIy 760
Db      2221 GTGACATCTCAGAGATGGAACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Qy      761 CysLeuSerAPLeuValIleAPProGIyAPASnAlaIthrAPValArgIyTyTyraIthr 780
Db      2281 TGCTATCTGATTTGGTTCATCCGAAACAAATGCTATGATGATGATGATGATGATGAT 2340
Qy      781 ArgTyGIyValPheTyPheSerIySerGIyTyTyraIleIleThrGIyLeuIySerValIleu 800
Db      2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGGAATACTGAAGGCTGTGTTCTTGA 2400
Qy      801 Ser 801
Db      2401 TCA 2403

```

```

; Sequence 3, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseeryoung, Katherine W.
; APPLICANT: Vitha, Stanislav
; APPLICANT: Kosharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: Use
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-3

Alignment Scores:
Pred. No.: 0 Length: 3667
Score: 3837.50 Matches: 800
Percent Similarity: 78.9% Conservative: 0
Best Local Similarity: 78.9% Mismatches: 1
Query Match: 94.4% Indels: 214
DB: 7 Gaps: 5

US-10-600-070b-2 (1-801) x US-10-600-070-3 (1-3667)
Qy      1 MetGIyAlaLeuSerHisValGIyIleGIyLeuSerProPheGIyLeuCyArgLeuPro 20
Db      481 ATGAGAGCTTGATGATCAGCGGATGGTCTCTCCCATTCATTAATGATGATGATGATGAT 540
Qy      21 ProIleThraIyLeuIyArgArgSerHisIleIthrSerThraIyIleCysSerAlaSer 40
Db      541 CCGGCGACCACAAAAGCTCCGACGTACCCACCAACACCTTACACATATGCTGCTGCTGCT 600
Qy      41 LyeThraIleAPAPArgLeuIySerAPAPheAPNheThraSerAPSerSerSerSer 60
Db      601 AAATGGGCGAGCGTCTTCTCTCCGATTCATTAATTAATTAATTAATTAATTAATTAATTA 660
Qy      61 PheAlaThraIleThraIyThraIyIleuValSerIyProProSerIleAPAPArgPro 80
Db      661 TTGCGACCGGCGACCAACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy      81 GluArgHisValProIleProIleAPPheTyTyraIleValIleuGIyAlaGIyIleIleIle 100
Db      721 GAAAGCGACGTCCCGATCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
Qy      101 LeuThraAPGIyIleArgArgAlaPheGIyAlaArgValSerIyAPProGIyAPNheGIy 120
Db      781 TTAACCGATGATTCAGAGAGATTCGAAAGCTTTCGAAACCGCGCAATTCGCT 840
Qy      121 PheSerAPAPAlaLeuIleSerArgArgGIyIleuGIyAlaIleuGIyAlaIleuGIy 140
Db      841 TTCAGGAGCAGAGCTTAAATCAAGCCGAGACAGATTCCTCAAGCTCTTGGCAACTCTG 900
Qy      141 SerAPNProArgSerArgArgGIyTyraGIyIleuGIyLeuIyAPAPGIyGIyAlaIthr 160
Db      901 TCTTAATCTCGCTCTGAAAGAGTCAATGAAGGCTTCTTGATGATGAAGAGCTTACA 960
Qy      161 ValIleThraAPValProThraAP
Db      961 GTATCATCATGATGCTCTTGGGATTAAGTAATTCATTCGAAATTAATTAATTCCTTC 1020
Qy      169 -----LyeValProGIyAla 173
Db      1021 GTTTTAATTTCAATGATGATTAAGAGAGCACTTTTATCTAGTGAAGGTTCTCGGGGC 1080
Qy      173 AluGIyValLeuGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIyGIy 193
Db      1081 TCTCTGTATTCAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

```

QY 193 uLeuLyGluArgLeuProLySerPheLyGlnAspValValLeuValMetAlaLeuAl 213
 DB 1141 GCTTAAGAGAGAGTGGCTTACGTTTAAAGCAAGAGTGGTTTAAAGTGGGCTTGC 1200
 QY 213 aPheLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGly 233
 DB 1201 GTTCTCGATGTCGAGGGAGTGCATGGCATTTGACCTCATTTTATTAATGAGTTA 1260
 QY 233 rGluPheValGluGluAlaLeuLyLeuLeuGln----- 244
 DB 1261 TGAGTTTGTGAGGAGACCTTGAAGCTTTTACAGTAGTTTGAAGCTTGTGTAATTTG 1320
 QY 244 ----- 244
 DB 1321 ACAGAGCTGGCTTATTAAGAACTTTCTGATTTGATTAAGCTTTGTAATTAAGCTTGTGTA 1380
 QY 245 -GluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuG 264
 DB 1381 GGAGGAAGAGCAAGTAGCCCTTGCACCGAATTTACGTGCACAAATTTGATAGAGCTTTGGA 1440
 QY 264 uGluIleThrProArgTyValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyVal 284
 DB 1441 AGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 1500
 QY 284 aAlaLyArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuThrSerValGlyGly 304
 DB 1501 TGGAAGAAAGCTAAATGTTAAAGCGGTGCGGAATATTTGTGCTGTGTGAGAGAG 1560
 QY 304 yGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlyLysPheMetAsnGluAlaPheLe 324
 DB 1561 TGAGAGATCAAGCTCTTGTGGGGTTTGACCCCGTGAAGAGTTTATGAATGAGCGGTTTT 1620
 QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
 DB 1621 ACAGATGACAGCTGCTGAGCAGGT-ATACAGATTAGATACCTTTTAAATTTCTTTAGC 1679
 QY 333 ----- 336
 DB 1680 ATGATATTAAGCTTAAAGCTTCTCATTTAAATGATGTTGTGTGAGGTGATCTTTTGTG 1739
 QY 336 AlaThrProSerAsnIleProAlaGluSerPheGluValTyArgValAlaLeuAlaLe 356
 DB 1740 AGCTACCCCAAGCAATTTCCAGCAGAGCTCATTTGAAGTTTACAAATTTGCACTTGCTCT 1799
 QY 356 uValAlaGlnAlaPheIleGlyLysLysProHleLeuLeuGlnAspAlaAspLysGlnP 376
 DB 1800 TGTGGCTCAAGCTTTTATGTTAAAGAGCCACCTTTTACAGGATGCTGATTAAGCAATT 1859
 QY 376 eGlnGlnLeuGlnGlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyArgAsp 396
 DB 1860 CCAAGCAATTCACACAGGCTAAGGTAATGCTATGGAATTCCTGCGATGTTGATGATAC 1919
 QY 396 rArgAsnAsnIleArgLysIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleG 416
 DB 1920 ACCGAATTAATGGAGAGTAACTTCGCTCTAGAAAGGAGACTCTGTCACTGCTTAAAGG 1979
 QY 416 yLysValAspGlyCyAlaArgMetIleLeuGlyLeuAspSerGlyLysAspSerGlnTyArgAs 436
 DB 1980 CAAGTTGATGAATGCGGTATGCTGGCTTAAAGACATGAGGATTCACAAATTAAGAA 2039
 QY 436 nProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAsnAspAspLeuP 456
 DB 2040 TCCAGCTAATGTGAGATTTTGTGGAATTCAAATCGTGAATGACAAATATGATCTCC 2099
 QY 456 oGlyLeuCyAlaLeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAs 476
 DB 2100 TGAACATATGCAATTTGTGAAACCTGCTTGGCAGGGGTTGTCTTTCTTACGTTCAAGA 2159
 QY 476 pThrLyAspLysLysPheLysLeuGlyAspTyTyTrpAspAspProMetValLeuSerTy 496
 DB 2160 CACCAAAAGATTAATAATTTAACTCGGGACTACTATGATGATCTCATGTGTTTGAAGTTA 2219

QY 496 rLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAlaIleThrMetAlaArg 516
 DB 2220 CTTGGAAAAGCTGAGAGTATTCAGGGTTCTCTTTAGCTGCTGCTCAATAGGCAAG 2279
 QY 516 gIleGlyAlaGluHleValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSe 536
 DB 2280 GATTGAGCCGAGCATGTGAAGACTAGTGCATAGCAGGCACTGCAGAAATTTTCTTCC 2339
 QY 536 rArgTyThrAspArgAsnSerAlaGluProLyAspValGluGluThrValPheSerVal 556
 DB 2340 CCGCTATACAGATGAAGAACTCGGCTGAACCAAGAGATGTGCAGAGACAGTGTATG 2399
 QY 556 lAspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAl 576
 DB 2400 AGATTCCTGTTGTATACATATAGCGCTGATGTGAGCCCTGCTTATTTGAGAAAGC 2459
 QY 576 aValArgProSerGlyLysAsnPheGluThrAsnAspTyValAlaArgAlaGlyValSerG 596
 DB 2460 TGTAAAGCCCTGAAACCTTGAACATAATGATTAATGCAATTCAGCTGGGGTCTCAGA 2519
 QY 596 uSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSe 616
 DB 2520 GAGTAGCGTGTATGAACATACCTGTGAATGTCGTTGCTGATATGTTAAAGAGGCAAG 2579
 QY 616 rValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTy 636
 DB 2580 TGTGAAGATCTTACCTGCTGCTGTGTGCAATTTGACATATTTCACTGTTCAAGCAAGTA 2639
 QY 636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGlySerAs 656
 DB 2640 TTTTCTTAAAGCAGCTCATCTTTTCAACCAAGATATGTTTCTTATGGAATCTGA 2699
 QY 656 pValAlaThrIle----- 660
 DB 2700 TGTGCTTACCATATGATATTAATGATCAATTTTCATATATCTGATTCACAAAT 2759
 QY 660 ----- 660
 DB 2760 ATGCTGTTTGTGAGCTAAGAAACATAGTTCCCATTAATATATCCCAAAAGTTGTAC 2819
 QY 660 ----- 660
 DB 2820 CAAGATTAACAAGTGTGAGTAATTTCACTAATATGCTGCTGAATTTTGTATCA 2879
 QY 660 ----- 660
 DB 2880 ACTGTAGACAGAAATGTAATTTCACTCACTCAATTTCTGTTAAGATTAAGATTA 2939
 QY 660 ----- 660
 DB 2940 GAGATTCCTTAAGTGTGCTTGTCCAACTTTTCTTGAATTTTCTTTTGTGAT 2999
 QY 661 --GlySerValArgAlaAspAspSerGlyLysLeuProArgMetAspAlaArgThrArg 680
 DB 3000 TAGGGTCAAGTCAAGAGCTGAGATTCAGAGACATTCACAAATGATGCTAAGAGCTGAG 3059
 QY 680 lAsnIleValSerTyTrpGlnLysIleLysSerLeuAlaPheGlyProAspHleArg 700
 DB 3060 AGAATATGATATCCAAATGCGCAGAAATTAAGTCTGCTTGTGGGCTGATCACCGCA 3119
 QY 700 lGluMetLeuProGlu----- 705
 DB 3120 TAGAAATGTTAACAGAGGTGAGGAAATTAATCTAATCAATCAATGATGTTGAAAACTG 3179
 QY 706 -----ValLeuAspGly 709
 DB 3180 TTGACATGATTAATAGCTGAGCTTGTGTTGATTCGTATTTAATTAAGGTTTGTGATGG 3239
 QY 710 ArgMetLeuLysIleThrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTy 729
 DB 3240 CGAATGCTGAAGATTTTGAAGCTGACAGAGCAGCTGAAGATCGCGAGCTTGGGTGTTAT 3299
 QY 730 AspTyThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArg 749

|||||
Db 3300 GATTATACACTGTTGAAGACTATCTGTGACAGTGTGACAGTCCAGACGATGAAACCGT 3359
750 AAlenValGluValThrLeuGluSerAlaCysLeuSerAspLeuValHisProGlu 769
Db 3360 GCTCTGTGTGAAGCACTGTGAGAGAGCTGTGCTGTCTATCTGATTTGATTCATCCAGAA 3419
770 AsnAsnAlaThrAspValArgThrTyrThrArgTyrGluValPheTyrSerLysSer 789
Db 3420 AACATCTCTACTGATGTCAAGACTTACACAAAGATACGAAGTTTCTGTCTCAAGTCA 3479
790 GATTrpValLeuThrGluGlySerValLeuAlaSer 801
Db 3480 GGGTGAAATCACTGAAGCTGTCTTCTTCATCA 3515
RESULT 7
US-10-600-070-10
; Sequence 10, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kokharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-10
Alignment Scores:
Pred. No.: 0 Length: 3667
Score: 3823.50 Matches: 798
Percent Similarity: 78.7% Conservative: 0
Best Local Similarity: 78.7% Mismatches: 3
Query Match: 94.1% Indels: 214
Gaps: 5
US-10-600-070b-2 (1-801) x US-10-600-070-10 (1-3667)
QY 1 MetGluAlaLeuSerHisValGlyLeuSerProPheGluLeuCybArgLeuPro 20
Db 481 ATGGAACTCTGAGTCACTGCGGCAATGGTCTCTCCCATTCCAATTATGCGCATACCA 540
QY 21 ProAlaThrThrIleValLeuArgSerHisAsnThrSerThrIleCysSerAlaSer 40
Db 541 CCGCGCAGCAAAAGCTCCGACGTGACCAACACCTCTACACTTCTGTCTGCGCAGC 600
QY 41 LysTrpAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
Db 601 AAATGGCCGACCTCTTCTCTCCGACTTCAATTTCCTCCGATTCCTCTCTCTCC 660
QY 61 PheAlaThrAlaThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
Db 661 TTGCGCACCGCCACCAACACGCGCACTCTGCTCTGCGCACCATCATTTGATCGTCC 720
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100
Db 721 GAACGCCACGCTCCCATCCCATTTGATTTTACACAGATATTAGAGCTCAAAACATTTCC 780
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
Db 781 TTAAACCATGATCAAGAGAGCAATTCGAAGCTTGGGTTTCGAACCGCGCATTTGGT 840
QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCysGluThrLeu 140

Db 841 TTCAGCGACGACGCTTTAATCAACCCGAGACAGATTTCTTCACTGCTCCGAACCTCTG 900
QY 141 SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
Db 901 TCTTAATCTCGCTCTGAAGAGAGTACATAGAAAGCTTCTTATATATGAAGAGCTACA 960
QY 161 ValIleThrAspValProTrpAsp-----LysValProGlyVal 168
Db 961 GTCATCACTGATGCTCTTGGATTAAGTAAATTCATTCGGAATATTAAGTTTCTTCC 1020
QY 169 -----LysValProGlyVal 173
Db 1021 GTTTTAATTTCAATGAATTCGATTAAGAGAGAACTTTATCTAGAGAGGTTCTCGGAGC 1080
QY 173 AluCybValLeuGluGluGlyValGluThrGluIleValLeuArgValGlyValAlaLeu 193
Db 1081 TCTCTGTGATTTCCAGAGAGAGTGTAGACTGAGATAGTTCTTCGGGTTGTAGAGCTCT 1140
QY 193 ULeuValGluArgLeuProLysSerPheLysGlnAspValValLeuValMetAlaLeuAl 213
Db 1141 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTGTTAGTTATGCGCTTGC 1200
QY 213 AsnLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyr 233
Db 1201 GTTCTCGATGTCTCGAGGAGTCTATGGCATTGATCCACCTGATTTTATTTACTGCTTA 1260
QY 233 rGluPheValGluGluAlaLeuLysLeuGln----- 244
Db 1261 TGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGTATGTTGACTTCTTGGTAAATTGG 1320
QY 244 ----- 244
Db 1321 ACAGCGTTGGCTTTAAGAACTTCTTGATTTGATTTGTTATGAGCTTGTGTA 1380
QY 245 -GluGluGluValAspSerLeuAlaProAspLeuAlaGlnIleAspGluThrLeuGlu 264
Db 1381 GGAGGAAGGAGCAAGTAGCTTGTGACCGATTTACGTGACCAAAATGATGAGACTTTGGA 1440
QY 264 UGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAspTyrAl 284
Db 1441 AGAGATCACTCCGCTTATGTCTTGAAGCTTACCTTGGCTTACCGCTTGTGATATATACG 1500
QY 284 aAlaValArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSerValGlyGly 304
Db 1501 TCGGAAAAGACTAAATGCTTTAAGCGGTGCGGAATATTTGTGCTGTGAGAGAGG 1560
QY 304 YGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLe 324
Db 1561 TGGAGCATCAGCTCTGTGTGGGGGTTTGACCCGTGAGAAATTATGAATGAGCGGTTTT 1620
QY 324 UArgMetThrAlaAlaGluGlnVal----- 332
Db 1621 ATGAATGACAGCTGTGAGAGAGGT-ATACAGTTTACATCTTTTAAATTTCTTTAGC 1679
QY 333 -----AspLeuPheVal 336
Db 1680 ATGATTAATCTTAGGTTTCTCATTTTAATGATGTTGTGTGAGGTTGATCTTTTGT 1739
QY 336 LAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLe 356
Db 1740 AGCTACCCCAAGCAAAATTCACAGCAGATCATTTGAAGATTGAGAGTTGACCTGTCT 1799
QY 356 ValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPh 376
Db 1800 TGTGGCTCAAGCTTTATGTGTAAGAGCCACCTTTTACAGAGATGCTATAGCAATT 1859
QY 376 eGlnGlnLeuGlnGlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspTh 396
Db 1860 CCAAGCAACTTCAACAGGCTTAAGGTAAATGCTTAAGGATTCCTGCGATGTGTATGATAC 1919
QY 396 rArgAsnAspTrpGluIleAspPheGlyLeuGluValGlyValLeuCybAlaLeuLeuIleG 416
Db 1920 ACCGAATTAATTTGGAGATGACTTCTGTCTAGAAAGGAGACTCTGTGCACTGCTTATAG 1979

```

Qy 416 YLysValAspGluCysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAs 436
Db 1980 CAAAGTTGATGAAATGCCGTAATGCTGGCTTGGAGCTTGAACAGTGAAGATTCAACAATATGAGAA 2039
Qy 436 nProAlaIleValGluPheValLeuGluIleuSerAsnArgAspAspAsnAspAspLeuP 456
Db 2040 TCCAGCTATGTGAGAGTTCTTTTGGAGAAATCGATGACAAATGATGATGATCTCCC 2099
Qy 456 OGlyLeuGlyLeuLeuLeuGluThrTrpLeuAlaGlyValAlaPheProArgPheArgAs 476
Db 2100 TGGACATATGCAAAATGTTGGAAACCTGGTGGCAGGGGGTGTCTTTCTTGCTTGACAGAA 2159
Qy 476 pThrLysAspLysLysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyr 496
Db 2160 CACCAAGATTAATAAATTATTAACCTGGAGAACTACTATGATGATCTATGGATTGAGATTA 2219
Qy 496 rLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaAr 516
Db 2220 CTGGAAAGAGTGAAGTACTTCAAGGTTCTCTTTAGCTGCTGCTGCACTATGGCAG 2279
Qy 516 gIleGlyAlaGluHisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSe 536
Db 2280 GATTGAGCGCAGACATGTGAACCTAGTCTATGCAAGGCATGCAGAAATTTTCTTCTTC 2339
Qy 536 rArgTyrThrArgAspAsnSerAlaGluProLysAspValGlnGluThrValPheSerVa 556
Db 2340 CCGCTATACAGATGAAGAACTCGGCTGAACCCAGAGATGTCAGAGACATGTTTAAAGT 2399
Qy 556 lAspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAl 576
Db 2400 AGATCTGTTGGTAACAATGTAGCCCTGATGTGTGAGCCCTGCTTTATTCAGAAAGC 2459
Qy 576 aValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGI 596
Db 2460 TGTAAACCTCTCGAAAACTTGTAAACTAATGATATGCAATTCGAGCTGGGCTCTCAGA 2519
Qy 596 uSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluLysLe 616
Db 2520 GAGTACCGTGAAGAACTACTGTGAAAGTCCGTTGCTGATGATGTTAAAGAGGCGAG 2579
Qy 616 rValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnTyr 636
Db 2580 TGTGAAGATCTTACGCTGCTGCTGCGCAATGGAATGATTCATCTGTTCAGCCAGAAAGTA 2639
Qy 636 rPheLeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
Db 2640 TTTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTCTTCTATGGAATCTGA 2699
Qy 656 pValAlaThrIle----- 660
Db 2700 TGTGCTACCATAGTATGATTAATGATGCAATTTTCATATATCTGCATTTGCTCAAAAT 2759
Qy 660 ----- 660
Db 2760 ATGCTTGTTTTGTGAGCTAAGAACATAGTTCCCACTTAATACATGTCCAAAGTTGTAC 2819
Qy 660 ----- 660
Db 2820 CAAGATTACAAGTTGCTGAGTAATTTTCACTAATTATGCTGCTGAATTTTGTATCA 2879
Qy 660 ----- 660
Db 2880 ACTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTGAATAACGTAGATTA 2939
Qy 660 ----- 660
Db 2940 GAGATTGCTTATGCTGTGCTTTGTCCAATTTTCTTCTTGAATTTTCTTTTGATTT 2999
Qy 661 --GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
Db 3000 TAGGGTCAGTCAAGAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGTAGTGAAGTCAG 3059

```

```

Qy 680 lAsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgI 700
Db 3060 AGAATATAGTATCCAAATGTCGACAGAAATTAAGTCTCTGCTTTTGGCCCTGATCACCGA 3119
Qy 700 lGluMetLeuProGlu----- 705
Db 3120 TAGAAATGTATCCAGAGGTGAGGAAATTAATCTACATTCATCAATCAATGTTGTGAAAACTG 3179
Qy 706 -----ValLeuAspGly 709
Db 3180 TTGACATGATTTATACCTGTGTCGCTTGTATTTGATTTCTGTTATTAAGTTTGGAAAGG 3239
Qy 710 ArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyr 729
Db 3240 CGAATCTGAAGTTTGGATGACAGACAGCTGAAATCTCGAGCTTGGGTTGTTAT 3299
Qy 730 AspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArg 749
Db 3300 GATTATACACTGTTGAACATATCTGTTGACAGTGTACAGTCTCAGCAGATGGAAACCGT 3359
Qy 750 AlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGlu 769
Db 3360 GCTGTGTGAGAACCACTCGAGAGAGTGTCTGTCTATCTGATTTGTTGATCCAGAA 3419
Qy 770 AsnAsnAlaThrAspValArgThrTyrThrArgTyrGluValPheTrpSerLysSer 789
Db 3420 AACATGCTACTGATGTGCAACCTTACACACAAAGTATGAAATTTTCTGTGTCCAAGTCA 3479
Qy 790 GlyTrpLysIleThrGluGlySerValLeuAlaSer 801
Db 3480 GGGTGAATAATCACTGAAGGCTCTGTTCTTGATCA 3515

```

```

RESULT 8
US-10-600-070-126
; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oeteryoung, Katherine W.
; APPLICANT: Vitina, Stanislav
; APPLICANT: Kosharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placoid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-600-070-126

```

```

Alignment Scores:
Pred. No.: 1,14e-168 Length: 2283
Score: 1775.50 Matches: 390
Percent Similarity: 62.5% Conservative: 119
Best Local Similarity: 47.9% Mismatches: 213
Query Match: 43.7% Indels: 93
DB: 7 Gaps: 16

```

```

US-10-600-070B-2 (1-801) x US-10-600-070-126 (1-2283)
Qy 12 SerProPheGlnLeuCysArgLeuProAlaThrThrLysLeuArgArgSerHisAsn 31
Db 40 GGGCATTTGCCCTTTCCTCCCTGCGCCGCGCGCGCGCGCGCGCGCGCGCGCT 99
Qy 32 ThrSerThrThrIleCys--SerAlaSerLysTrpAlaAspArgLeuLeuSerAspPhe 50
Db 100 CACCCCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCGCATTC 159
Qy 51 AsnPhe-----ThrSerAspSerSerSerSerSerPheAlaThrAlaThr 65

```

Db 160 CACCTCTCCCAACGCGCGCCCTCCGACCCGCGCTCCCGCGCCCGCGCGCC 219
::: ||||| :::
Qy 66 ThrThrAlaThrLeuValSerLeuProProSerIleAspArgProGluArgHisValPro 85
::: ||||| :::
Db 220 GCGCCCTCCGCTCCCTCTTGCCTCTTCCCGCAGCGCCGCAAGCTCCCTCCG 279
::: ||||| :::
Qy 86 IleProIleAspPheThyArgValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIle 105
::: ||||| :::
Db 280 CTCCAAGTCGATTTCTTCAAGAGTTCTAGGGGCGAGACCACTTCTCTGGCGATGGCATC 339
::: ||||| :::
Qy 106 ArgArgAlaPheGluAlaArgValSerLysProProGlnPheGlyPheSerAspAspAla 125
::: ||||| :::
Db 340 AGAGAGCGCTTCAAGGACGAGATAGCCAGCCAGCATGATGCTACAGCAGATGCT 399
::: ||||| :::
Qy 126 LeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGlnThrLeuSerLeuProArgSer 145
::: ||||| :::
Db 400 CTGTGTGCTGTGACAAATGCTGCAGATTGCCATGACCTCTCATGAAACGAACCTCC 459
::: ||||| :::
Qy 146 ArgArgGluTyArgGlnGlyLeuLeuAspAspGluGluAlaThrValIleThrAspVal 165
::: ||||| :::
Db 460 CGCAGCTCAGTATCATGTGCGCTTTCTGAGAACCGTGAAGAACTCTCACCATGATATT 519
::: ||||| :::
Qy 166 ProTrpAspLysValProGlyAlaLeuCysValLeuGlnGlyGlyGlnThrGluIle 185
::: ||||| :::
Db 520 GCTGGGACAG-----GAGGCTGGGAGGACCTTGCT 552
::: ||||| :::
Qy 186 ValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLysSerPheLysGlnAsp 205
::: ||||| :::
Db 553 GTGCTTGTAACTGGAGAAAGCTTCTGATCGGCGCACCCAGGCGCTTCAAGCAGAGAC 612
::: ||||| :::
Qy 206 ValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAsp 225
::: ||||| :::
Db 613 GTGGTCTACCGCATGGCTCTGCTTATGTGATCTATCAAGGATGCTATGGACAGAC 672
::: ||||| :::
Qy 226 ProProAspPheIleThrGlyTyArgLysPheValGluGluAlaLeuLysLeuGlnGlu 245
::: ||||| :::
Db 673 CTCCTCAGATGATATGGCTCTGCGAGTGTGCGAGGCGCTCTCAAGCTCTTGAGAA 722
::: ||||| :::
Qy 246 GluGlyAlaAspSerLeuAlaProAspLeuArgAlaGlnIleAspGlnThrLeuGlnGlu 265
::: ||||| :::
Db 733 GATGAGCAGCAAGCATCTCGCACCTGATCTGCTTCACAGTATGATGAACCTTCGAGAG 792
::: ||||| :::
Qy 266 IleThrProArgTyValLeuGlnLeuLeuGlyLeuProLeuGlyAspAspTyAlaAla 285
::: ||||| :::
Db 793 ATTACCTCTGCTGTGATTTGAGAGCTTCTCTCCCTTCTATGACACAGCATCATTAAG 852
::: ||||| :::
Qy 286 LysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGly 305
::: ||||| :::
Db 853 AAGCGCAAGAAAGGGCTTCAAGGTGCGAAGAAACATTTTGTGAGCGTTGCGAGAGGT 912
::: ||||| :::
Qy 306 AlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetLeuGluAlaPheLeuArg 325
::: ||||| :::
Db 913 ATTTGACCTCTGAGAGGATTTTCTGTGAAGCCCTTCATGAACGAGGCTTTTGTAGG 972
::: ||||| :::
Qy 326 MetThrAlaIleGlnGlnValAspLeuPheValAlaThrProSerAsnIleProAlaGlu 345
::: ||||| :::
Db 973 ATGACATCAATTACAGATGATTTCTTTCAAAAACCCAGATAGCATTTCTCTCGAA 1032
::: ||||| :::
Qy 346 SerPheGluValTyArgLysValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLys 365
::: ||||| :::
Db 1033 TGGTTTGAATTTTACAATGATAGCACTTGCAATGTGCTCAAGCAATTATTAAGTAAAG 1092
::: ||||| :::
Qy 366 ProIleLeuLeuGlnAspValAspLysGlnPheGlnGlnLeuGlnGlnAlaLysValMet 385
::: ||||| :::
Db 1093 CCACAAATTCATCATGATGCGGAGTGAATCTTTTGAACACTCCAGAACTTCAACATTTG 1152
::: ||||| :::
Qy 386 AlaMetGluIleProIleMetLeuTyArgPheThrArgAsnAsnTrpGluIleAspPheGly 405
::: ||||| :::
Db 1153 TCT-----CATTATGCTTATGATTAAT-----GAGATGAGACTTGCA 1188
::: ||||| :::
Qy 406 LeuGluArgGlyLeuCysAlaLeuLeuIleGlyLysValAspGluCysArgMetTrpLeu 425
::: ||||| :::

Db 1189 TTGAAAGGCACTTCTGCTCATTTCTAGTCGAGATGTTTACGAATGCGAAATGCGCTT 1248
::: ||||| :::
Qy 426 GlyLeuAspSerGluAspSerGlnTyArgAsnProAlaIleValGluPheValLeuGlu 445
::: ||||| :::
Db 1249 GGAATTTGAATGAAGCTTACCATACAGAGACCCCAAAATTTCTAGATTATTTGAGACC 1308
::: ||||| :::
Qy 446 AsnSerAsn--ArgAspAspAsnAspAspLeuProGlyLeuCysAlaLeuLeuGlnThr 464
::: ||||| :::
Db 1309 AACTACATCATCAGTGAAGATGATCTTCTTCAGGCGCTGTGCAAGCTTTTGAGACT 1368
::: ||||| :::
Qy 465 TrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLysPheLysLeu 484
::: ||||| :::
Db 1369 TGGCTTATCTTTGAGGTTTTTCTTACAGACAGATATCTCGGGCATGCAAGTTCAACTT 1428
::: ||||| :::
Qy 485 GlyAspTyTyArgAspAspProMetValLeuSerTyLeuGluArgValGluValGln 504
::: ||||| :::
Db 1429 GGAAGATTTACTAGATGATCAGAAATTTTAACTACTCTGAAAGATGAGTGGGTGTGT 1488
::: ||||| :::
Qy 505 GlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyValAlaGlu----- 520
::: ||||| :::
Db 1489 GCTTCTCATTTGGCTGCTGCTGCTGCTATGCAAACTTGCTCAACCTCAAGCTGCA 1548
::: ||||| :::
Qy 521 -----HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPhePro--SerArg 537
::: ||||| :::
Db 1549 CTGTGACTGTGAATCAAAATGCTATTCAGGCTTCAACAGGTTTTTCAATTGATAGAA 1608
::: ||||| :::
Qy 538 TyTrpAspArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAsp 557
::: ||||| :::
Db 1609 CAGTTAGACAGTGCACCATGGAATAACT----- 1638
::: ||||| :::
Qy 558 ProValGlyAsnAsnValGlyLysArgGlyGluProGlyValPheIleAlaGluAlaVal 577
::: ||||| :::
Db 1639 -----AAAGATGCG--CTTGCGGAGATATCTT----- 1662
::: ||||| :::
Qy 578 ArgProSerGluAsnPheGlnThrAsnAspTyAlaIleArgAlaGlyValSerGluSer 597
::: ||||| :::
Db 1663 -----GAAATTTTACACG----- 1677
::: ||||| :::
Qy 598 SerValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerVal 617
::: ||||| :::
Db 1678 -----GAAATGCACTGCTCATGATTCGAAGAAATGCGCTTG 1716
::: ||||| :::
Qy 618 LysIleLeuAlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyPhe 637
::: ||||| :::
Db 1717 AAGATTATCTGCTGCGCCACACTGTGTGCACTGTGCGAGTATTTGGCGCCAAATAT-- 1773
::: ||||| :::
Qy 638 LeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAsp--- 656
::: ||||| :::
Db 1774 -----TTGCTCTGTAAGAGGCCCTTTCTGCTATTGAGATGAGCAT 1815
::: ||||| :::
Qy 657 -----ValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeu----- 671
::: ||||| :::
Db 1816 GGAATCTGTGGCACTTCTTAATATGTGCACTGATCAGATGATCTGTCACATGATGAAGAT 1875
::: ||||| :::
Qy 672 -----ProArgMetAspAlaArgThrAlaGluAsnIleValSerLysTrpGln 687
::: ||||| :::
Db 1876 CCAGTACATATTTCTGAATGATGATGCGAAGCTGCGCAAGATATTTGTCGCAAGTGCAG 1935
::: ||||| :::
Qy 688 LysIleLysSerLeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeu 707
::: ||||| :::
Db 1936 AGTATCAAACTTAAGGCTTTGGACCAAGAACATTCGTCATCATTTGCAAGAGTTCTT 1995
::: ||||| :::
Qy 708 AspGlyArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeu 727
::: ||||| :::
Db 1996 GATGGCAATGCTTAAGGTGATGATGCAAGCCAGACGCGAGATTAAGGTGATGGGTGG 2055
::: ||||| :::
Qy 728 ValTyAspTyTyThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGly 747
::: ||||| :::
Db 2056 TTCTGGAGATATCACTATCCGATGATGACGATGATGATGATCTCTCTTCAAGTATGTG 2115
::: ||||| :::
Qy 748 ThrArgAlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHis 767
::: ||||| :::
Db 2116 CGACGAGCGACTGTGAGGCTAGATTTGATGAGGCGAGGCCCACTTACTGATTTACTGAG 2175
::: ||||| :::


```

QY 768 ProGluAspAlaThrAspValArgThrThrThrArgTyrGluValPheTrpSer 787
DB 2176 CCCAGAAACAAATGATTATATGACAAATAACCTACCGGATGAGATGCGCTTCTCC 2235
QY 788 Lys---SerGlyTyrPylIethrGluGlySerValIleuAlaSer 801
DB 2236 AACCTAGAGGCTGAGATGACGAGAGGAGCTCTCAAGTCG 2280

RESULT 9
US-10-424-599-129007
; Sequence 129007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129007
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87499C.1
US-10-424-599-129007

```

```

Alignment Scores:
Pred. No.: 3,27e-104 Length: 1146
Score: 1134.00 Matches: 223
Percent Similarity: 80.7% Conservative: 48
Best Local Similarity: 66.4% Mismatches: 53
Query Match: 27.9% Indels: 12
DB: Gaps: 3

```

US-10-600-070B-2 (1-801) x US-10-424-599-129007 (1-1146)

```

QY 16 LeuGlyArgLeuProAlaThrThr-----LysLeuArgArg 28
DB 148 CTTGCACTCCGACCCACCACTCACTTCACTAACTTAACCACTCTCCGT 207
QY 29 Ser-----HisAnthSerThrThrIleCysSerAlaSerIleTyrAlaAspArgLeu 46
DB 208 TCTCCTCCTAGCGCGCGCGCTTCTTATCGCCACCAATGCGGAGCGCTC 267
QY 47 LeuSerAspPheAsnPhetTrpSerAspSerSerSerSerPheAlaThrAlaThrThr 66
DB 268 ATGCGCATTTCCAAATCTCGCGAGCGCGCGCTCC-----ACCTTCACCTCC 318
QY 67 ThrAlaThrIleValSerLeuProSerSerIleAspArgProGluArgHisValProIle 86
DB 319 ACTCTGAGCCCTCTCTCGCTCCCTCCCGCTCGATCTCTGAGCGCTAGCGTGCATC 378
QY 87 ProIleAspPheTyrGluValIleGluGlyAlaGlnThrHisPheLeuThrAspGlyIleArg 106
DB 379 CCGCTCGACCTGACCGCAATCTCGGCGCGGAGCGCACTCTCTCGCGCGAGCATTCGC 438
QY 107 ArgAlaPheGluAlaArgValSerIleProProGluPheGlyPheSerAspAlaLeu 126
DB 439 AGGCGCTACAGCGCAAGTCTCGAAGCTCTCAGACCCCTTCAGCAAGCGCAAGCGCTTCA 498
QY 127 IleSerArgArgGlnIleGluGlnAlaCysGluThrIleuSerAsnProArgSerArg 146
DB 499 ATAGCGCGCGCAATCTCTCGAAGCGCTCGGAGAACTTACGTATCTTACTTCACAGA 558
QY 147 ArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThrValIleThrAspValPro 166
DB 559 AGAGATACATCAAGACCTTGTGACGAGAGAGAGCGCCATTCCTCACTCAATTCCT 618

```

```

QY 167 TrpAspLysValProGluAlaLeuCysValIleGluGlnGlyGlyGluThrGluIleVal 186
DB 619 TTGCAACAATGTTCTTGAGCGCTTGTGTGTGTGACAGAACTGAGAGACGACCTGTG 678
QY 187 LeuArgValGlyGluAlaLeuLeuGlyGluArgLeuProLysSerPheLysGlnAspVal 206
DB 679 CTGAGATTTGGGAGGGTTTGTCTTGAAGAGAGGTTCGCAAGACGTTTAAGACGAGATGTT 738
QY 207 ValIleValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspPro 226
DB 739 GTGTGGCTATGGACACCGCATTTGTGAGGTGTGCAAGGAGATGCTATGAGCTGTGTCCA 798
QY 227 ProAspPheIleThrGlyTyrGluPheValGluGluAlaLeuLysLeuGlnGluGlu 246
DB 799 CCGGATTCATTTGGCGCTGTGAGATGCTGAGAGGGCGCTTGAAGCTTTTGACAGAGAA 858
QY 247 GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGluIle 266
DB 859 GGGGCAACAGCGCTAGCTCAGATTTTACAGCAAAATGATGAGACGCTTAGAAGAGATA 918
QY 267 ThrProArgTyrValIleGluLeuLeuGlyLeuProLeuGlyAspAspTyrAlaAlaLys 286
DB 919 ACCCCAGTTGTGTGTGTGGAACCTTTAGCCCTGCTTGATGACGAACATCGAGCGCG 978
QY 287 ArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGlyValAla 306
DB 979 AGGAGAGAGGCTCTCTTGTGTGTGTCGTAACATTTGTGGCGGCTGTGAGAGGGGTGCA 1038
QY 307 SerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLeuArgMet 326
DB 1039 GCACCAATGCGCGGGGTTTCAACCGTGAAGACTTATGATAGGATCTTACACATG 1098
QY 327 ThrAlaAlaGluGlnValAspLeuPheValAlaThrProSerAsnIle 342
DB 1099 ACAGCGGCTGGAACAGGTGAACTTTGTGACCAACCAAGTACTATTT 1146

```

RESULT 10

```

US-10-425-115-81853
; Sequence 81853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81853
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174665C.1
US-10-425-115-81853

```

```

Alignment Scores:
Pred. No.: 8,53e-88 Length: 1411
Score: 973.00 Matches: 198
Percent Similarity: 69.8% Conservative: 54
Best Local Similarity: 54.8% Mismatches: 79
Query Match: 23.9% Indels: 30
DB: Gaps: 4

```

US-10-600-070B-2 (1-801) x US-10-425-115-81853 (1-1411)

```

QY 12 SerProPheGlnLeuCysArgLeuProProAlaThrThrLysLeuArgArgSerHisAsn 31
DB 152 GCGCGCTTGGCTTCTCAGCGCTGCGC-----CTCCGAGAGAGCCACAG 196

```



```

32 Thr-----SerThrThrIleCySerAlaSerIleThrAlaAspArgLeu 47
197 CCGCCACCGCCCTGGCCCTCCACCTCCGCCGCCGACCGCTCCCTTC 256
48 SerAspPheAsnPhe-----ThrSerAspSerSer 57
257 GCGGACTTCACACTCTCCCGCCGCCGACCCCGCAGCGCGCTCTCTCC 316
58 SerSerSerPheAlaThrAlaThrThrAlaThrLeuValSerIleProSerIle 77
317 TCGTCCCGCTTC-----GTCGCGACTTCCCC 343
78 AspArgProGluArgIleValProIleProIleAspPheTyrglnValLeuGlnIle 97
344 GAAGCCCGGACCGCCCTGCCCTCCCGGTGACTCTTCAAGATCTTGTCGGAG 403
98 ThrIlePheLeuThrAspGlyIleArgArgAlaPheGlnIleAlaArgValSerIleProPro 117
404 CCACATTTCTAGCGATGATTCGAGAGCGCTTGAGTCCGCGATAGCTAACCACT 463
118 GlnPheGlyPheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnIleAlaArg 137
464 CAGTATGGGTACGACACAGAACTCTTGTGGCGACGCGAAATGCTGCAGATTCGCCAT 523
138 GlnThrLeuSerAsnProArgSerArgGlnIleTyraAsnGlnIleLeuAspAspGlu 157
524 GATCTCTTCACAAACAGAGCTCGCCACCGATACACCGTCCTCCGAGACCT 583
158 GlnAlaThrValIleThrAspValProTyrAspIleValProGlyAlaLeuCyValLeu 177
584 GATGCGGACTCAACATGATGTTCCTGGATAGAGTTCAAGTGTGCTGTGCTTC 643
178 GlnGlnGlyIleGlnIleValLeuArgValGlnIleAlaLeuLeuGlnIleArg 197
644 CAGAGAGCTGGGAGGACAACTGTGTAGCACTGGAAGACACTTCTTCAGAGACCT 703
198 LeuProIleSerPheIleGlnAspValIleLeuValMetAlaLeuAlaPheLeuAspVal 217
704 CCACCTTAAGCGGTTCAAGCAGATGTGTCTTCAAGTGCATTTGCTTATGTGACATA 763
218 SerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGlnPheValGln 237
764 TCAAGGAGTCTATGGCAGACCCCTCCAGATGTAATCTGTGTGTAGAGCTTGAG 823
238 GlnAlaLeuIleLeuGlnIleGlnIleGlnIleValSerSerIleAlaProAspLeuArgAla 257
824 AGGCACTGAAAGCTCTCGCAGAGAGATGGCGCAAGCAATCTTGCACCTGAACCTTCA 883
258 GlnIleAspGlnIleLeuGlnIleIleThrProArgTyrValLeuGlnIleLeuGlnIle 277
884 CAAATTTGATTAACCTTTGAGAGATTAACCTCTGTGTATTTGAGCTCTTCCCTT 943
278 ProLeuGlyAspAspTyrAlaAlaIleArgLeuAsnGlyLeuSerGlyValArgAsnIle 297
944 CCACTAGATGAATAAATAAATAAATTAACGCCAAGAGCTCGCAAGGTGCAATAAATA 1003
298 LeuTyrSerValGlyIleGlyIleValAlaSerAlaLeuValGlyIleLeuThrArgGlnIle 317
1004 TTGTGAGTGTGGCAGAGGTGTGTATTTGCTAGTGTGAGAGAGATTTCTCTGAGAGCC 1063
318 PheMetAsnGlnIlePheLeuArgMetThrAlaAlaGlnIleAlaPheLeuPheValAla 337
1064 TTATGATGATAGGCTTTCTTGCAGATACATCACTAGCAGATGAGATTTCTTCTCTAAA 1123
338 ThrProSerAsnIleProAlaGlnIlePheGlnIleValTyrGlnIleAlaLeuVal 357
1124 ACACCGAATGATACCACTGAATGTGTGATGATCATATGTGTGCGACTTGGCCACATT 1183
358 Ala 358
1184 GCT 1186

```

```

RESULT 11
US-10-437-963-69933/c
; Sequence 69933, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; LENGTH: 1703
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70554C.1
US-10-437-963-69933
Alignment Scores:
Prd. No.: 2, 32e-86 Length: 1703
Score: 960.00 Matches: 227
Percent Similarity: 56.5% Conservative: 79
Best Local Similarity: 41.9% Mismatches: 147
Query Match: 23.6% Indels: 89
DB: 7 Gaps: 15
US-10-600-070b-2 (1-801) x US-10-437-963-69933 (1-1703)
286 LysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSerValGlyIleGly 305
1695 AGCGCCAGAAAGAGCGCTTCAAGGTGCGAGAAACATTTGTGAGCTTGCGCAGAGAGGT 1636
306 AlaSerAlaLeuValGlyIleLeuThrArgGlnIlePheMetAsnGlnIlePheLeuArg 325
1635 ATTGCTACCGCTTGGAGAGAGATTTCTCTGTGAAGCCTTATAGAACAGACTTTTGTAGG 1576
326 MetThrAlaAlaGlnIleValAspLeuPheValAlaThrProSerAsnIleProAlaGln 345
1575 ATGACATCAATTGAAACAGATGATTTCTTTCAAAAACACCAATAGCATTCCTCTGAA 1516
346 SerPheGlnIleValTyrGlnIleValAlaLeuValAlaGlnIlePheIleGlyIleVal 365
1515 TGGTTGAAATTTAACAATGATGACATTTGCAATGCTCGTCAAGCAATTAAGTAAAGG 1456
366 ProIleLeuLeuGlnIleAspAlaAspIleGlnPheGlnIleLeuGlnIleAlaValMet 385
1455 CCACAATTCATCAATGATGGCGATGATCTTTTGAACAATCCAGAGTTCAACATAGGT 1396
386 AlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAsnTyrGlnIleAspPheGly 405
1395 TCT-----CATTAAGCTTATGATTA-----GAGATGACCTTGCA 1360
406 LeuGlnIleGlyIleLeuCyValAlaLeuLeuIleGlyIleValAlaAspGlnIleCyArgMetTyrPleu 425
1359 TTGGAAGAGGCAATTCCTCATTTGATGATGAGAGATGTTAGCAAGTGCAGAAATGGGCTT 1300
426 GlyLeuAspSerGlnIleAspSerGlnIleTyrArgAsnProAlaIleValGlnIlePheValLeuGln 445
1299 GGAATTTGATTAATGATCTTCAACATACAGAGACCGCGCAATTTAAGTTATGTAGACC 1240
446 AsnSerAsn---ArgAspAspAsnAspAspLeuProGlyIleuCyValLeuLeuGlnIleThr 464
1239 AACTTAGCATCACTGATGAGAGATGATTTCTTTCAGAGGCTGTGCAAGCTTTTGGAGACT 1180

```



```

Db      525  ACTGTAATCAAGTCTCTTCAGGCTTCAGTAAGTTTCCCA-----569
Qy      541  ArgAsnSerAlaGluProIyAspValGlnGluThrValPheSerValaAspProValGly 560
Db      570  -----TTGATAGACAGCTGACATGTCAGGC 596
Qy      561  AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
Db      597  AAGGATATC-----CCAGT 611
Qy      581  GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValaAsp 600
Db      612  GATGATCTTGATTAATCT-----CTTGAAAACCTCGCCCAAGAGAGTGTCT 659
Qy      601  GluThrThrValGluMetSerValAlaAspMetLeuGluAlaSerValIleLeu 620
Db      660  GGAGACGCTATCCATGATTCAGAAATGACGCCCTG-----AGATTATC 704
Qy      621  AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleTyrPheLeuIleSer 640
Db      705  TCTGCTGGTACACTGCTTGCACTATTGTCAGTGTAGGTCTCAAGTCTGCTGCTGAC 764
Qy      641  SerSerSerPheGlnArgIleAspMetValSerSerMetGluSerAspValAlaIle 660
Db      765  AAGTCA-----CTTCCTGCTTATGGGCGCAGATGAGGTCTGTG 803
Qy      661  GlySerValArgAlaAspAspSerGluAla-----LeuProArg 673
Db      804  GCGGTGTTGACCTCATTTGATGTGTCAGCGCAGATGAGAGCCACTGAAATCCCTAGG 863
Qy      674  MetAspAlaArgThrIleGluAsnIleValSerIleTyrGlnIleIleValSerLeuAla 693
Db      864  ATGACGCGAAGTTGGCTGAAGACATTGTCGACAGGTGGCAAGCATCAAGTCCAGGCT 923
Qy      694  PheGlyProAspHisArgIleGluMetLeuProGluValIleAspGlyArgMetLeuIle 713
Db      924  TTGGGGCGCAACACACTGTCACGCGCATTCAGAAAGTCTCGCGCGCAACATGCTGAG 963
Qy      714  IleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValIleTyrAspTyrThrLeu 733
Db      984  GTGTGGACGACCGACCGCAGATGATGACCGACCGCTGTCTGCTGAGTACGCCCTC 1043
Qy      734  LeuIleLeuSerValaAspSerValThrValSerAlaAspGlyThrArgAlaLeuValGlu 753
Db      1044  TCCGGCGCTGACGATGACAGATCACTCTCCGTGACCGCGCGCGCGACGCGGTGAG 1103
Qy      754  AlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThr 773
Db      1104  GCCACGATCGAGAGAGTGGCGCGCTTACGACGTGACCGACCTTAAGAACGACAGCGCG 1163
Qy      774  AspValArgThrTyrThrArgTyrGluValPheTrpSerIle---SerGlyTrpIle 792
Db      1164  TACGACACCAAGTACACCGCCGCTACGAGTGACTCACTCCAGCGCAGCAGCGGTGAGG 1223
Qy      793  IleThrGluGlySerValLeuAlaSer 801
Db      1224  ATCACCGAAGCGCAGTCTCAAGTCTG 1250

RESULT 13
US-10-600-070-132/c
; Sequence 132, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20

```

```

; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 132
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (127)..(127)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (520)..(520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-132

Alignment Scores:
Pred. No.:      2,93e-69      Length:      561
Score:          784.00      Matches:      157
Percent Similarity: 98.1%      Conservative: 0
Best Local Similarity: 98.1%      Mismatches: 3
Query Match:      19.3%      Indels:      0
DB:              7          Gaps:      0

US-10-600-070b-2 (1-801) x US-10-600-070-132 (1-561)

Qy      642  SerSerPheGlnArgIleAspMetValSerSerMetGluSerAspValAlaThrIleGly 661
Db      559  TCATCTTTTCAACGACGAAGATATGTTCTTCTATGAAATCATGATCTGCTACATAGGG 500
Qy      662  SerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGluAsn 681
Db      499  TCAGTCAGAGCTGACCATTCAGAAACATTCCTCCAGATGATGCTAGACTGACGAGAAAT 440
Qy      682  IleValSerIleTyrGlnIleIleValSerLeuAlaPheGlyProAspHisArgIleGlu 701
Db      439  ATAGTATCCAGAGTGGCGAAGATTAAGTCTCTGCTTGGGCTGATCACCGCATAGAA 380
Qy      702  MetLeuProGluValIleAspGlyArgMetLeuIleTrpThrAspArgAlaAlaGlu 721
Db      379  ATGTACAGAGGTTTGGATGGCGAATGCTGAAGATTGGAAGTCTGACAGACGCTGAA 320
Qy      722  ThrAlaGlnLeuGlyLeuValIleTyrAspTyrThrLeuIleLeuSerValaAspSerVal 741
Db      319  ACTGCCAGCTTGGGTTGATTAAGTATTAACCTGTGAACATTAATCTGTTGACAGTGTG 260
Qy      742  ThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAlaCys 761
Db      259  ACGCTTCACAGATGAAACCCGTCTCTGTGTGAAGCACTTGAAGAGTCTGTCTGT 200
Qy      762  LeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrArg 781
Db      199  CTATCGATTGTTGTTATTCAGAAACAAATGACTGATGCTGAGAACTTACACAAACAGA 140
Qy      782  TyrGluValPheTrpSerIleSerGlyTrpIleIleThrGluGlySerValLeuAlaSer 801
Db      139  TACGAAGTTTCNGTCCAGTCAAGTGAAGTGAATAATCACTGAAGCTCTGTTCTTGATCA 80

RESULT 14
US-10-600-070-184
; Sequence 184, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of

```

```

; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600.070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 184
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Prunus persica
; NAME/KEY: misc_feature
; LOCATION: (21)-(21)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-184

```

```

Alignment Scores:
Pred. No.: 4,46e-64 Length: 631
Score: 733.50 Matches: 149
Percent Similarity: 81.0% Conservative: 21
Best Local Similarity: 71.0% Mismatches: 35
Query Match: 18.1% Indels: 5
Gaps: 3

```

US-10-600-070b-2 (1-801) x US-10-600-070-184 (1-631)

```

Qy 306 A l s e r a l e u a l g l y l y l e u t h r a r g l u y s p h e m e t a n g l u a l a p h e l e u 324
Db 1 G C A G T T G C A A T G C T G G G G G A T T C A C T A C G T G A A A A T T C A G A A C G A G C C T T C T T G 60
Qy 325 A r g m e t h r a l a g l u g l u a l a s p l e u p h e v a l a t h r p r o s e r a n i l e p r o l a 344
Db 61 C A R A T G C T C A G T G A G C A G G T T G A T T A T T T C A C T A C C C C A T A T A T C C C G C A 120
Qy 345 G l u s e r p h e g l u a l y r g l u a l a l e u a l a e u a l a g l u a l a p h e i l e g l y l y s 364
Db 121 G A A G C T T G A A G T T A T G G G G T G C T T G C G C T T G C T C A A C C T T G T T G T G A A 180
Qy 365 L y s p r o h i s t e u e n g l u a n e p a l a s p l y e g l i n p h e g l u n g l u n g l u a l y s v a l 384
Db 181 A A A C C T A T A C A T T C A A G A T G C G A A A C C T A T T C A G A A A C T T C A G C A G T T A A G G T A 240
Qy 385 M e t a l a m e t g l u i l e p r o l a m e l e u t y r a s p t h r a r g a n a n t r p g l u i l e a s p p h e 404
Db 241 A C A G C T C T A G A C A T T C T C T T G A C A C T A T A T A C A A A G A A C A G A C A G A T A G A C T T T 300
Qy 405 G l y l e u g l u a r g l y l e u c y a l a l e u l e u l e g l y l y s v a l a s p g l u c y a r g m e t t r p 424
Db 301 G C T T T G A G A G G G A C T G T T C A C T T C T T A G G G A C C T T G A T G A C A G T C G T T G C G T G 360
Qy 425 L e u g l y l e u a s p e r g l u a s p e r g l u t y r a r g a n p r o l a i l e v a l g l u p h e v a l l e u 444
Db 361 T T G G G C C T A G A C G A T A T G A T T C A C C A T A T A G A A A T C A C T G T T G A G C T T T G C T T G 420
Qy 445 G l u a n s e r - - - A s n a r g a s p a s p a n a s p - - - - - A s p l e u p r o g l y l e u c y l y s 460
Db 421 G A A A C T C A A A G A T G A C G A T G A C A A T G A C A A T G A C A T A T C T T C T G A C T T T G A A G 480
Qy 461 L e u l e u g l u t h r t r i p l e u a l a g l y v a l p h e p r o a r g p h e a r g a s p t h r l y s a s p l y s 480
Db 481 C T A T T G A A G C G T G T T G A T G A G G T G T A T T C C C A G G T T T A G A G A C A C A A A G A C A T A 540
Qy 481 L y s p h e l y l e u g l u a s p l y t y r a s p a s p p r o m e t a l l e u s e r t y l e u g l u a r g v a l 500
Db 541 G A G T T A G A C T G G A G C T A C T A T G A T G A T C T T A G A T C T T G A G A T A C T T A G A A A G G C T G 600
Qy 501 G l u a l a l a g l i n g l y s e r p r o l e u a l a l a 510
Db 601 G A T G C A C T A A T G A T T C A C C C T T A G C T G C T 630

```

RESULT 15
US-10-600-070-135
; Sequence 135, Application US/10600070
; Publication No. US20040139500A1

```

; GENERAL INFORMATION:
; APPLICANT: Osteeryoung, Katherine W.
; APPLICANT: Vilha, Stanislav
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600.070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
US-10-600-070-135

```

```

Alignment Scores:
Pred. No.: 1,76e-62 Length: 660
Score: 718.00 Matches: 141
Percent Similarity: 79.5% Conservative: 34
Best Local Similarity: 64.1% Mismatches: 35
Query Match: 17.7% Indels: 10
Gaps: 4

```

US-10-600-070b-2 (1-801) x US-10-600-070-135 (1-660)

```

Qy 25 L y s l e u a r g s e r h i s a n t h r s e r t h r l e c y s e r a l a s e r l y s t r p a l a s p 44
Db 15 A A A C T A A C C G T C C A T - - - - - T C T C G C G C G T C C G C A C C A G T A A T G G G C G A G 68
Qy 45 A r g l e u s e r a s p p h e a n p h e t h r s e r a s p e r s e r s e r s e r p h e a l a t h r a l a 64
Db 69 C G A C T A T T C C G A T T T C C A T T C C T C G G C G A C C T C C T C T C C A C C A C C A C C 128
Qy 65 T h r t h r a l a t h r l e u a l - - - - - S e r l e u p r o p r o s e r l l e a s p a r g p r o g l u a r g h i s 83
Db 129 T C G C C A C G T C A C T C T C A C T C T T C T T A C C T C T C C G A - - - - - G A A C G C A C 179
Qy 84 V a l p r o l e p r o l l e a s p p h e t y r g l u a l l e u g l y a l a g l i n t h r i s p h e l e u t h r a s p 103
Db 180 G T C A C T C C C T C T C A C C G T A C A A A A T C T C G G C G C G A A A G C A T T T T C G G T G A T 239
Qy 104 G l y l l e a r g a r g l a p h e g l u a l a r g v a l s e r l y s p r o p r o g l i n p h e g l y p h e s e r a s p 123
Db 240 G G T A T T C G A G A C T T A T G A A G G A A A T T C T C G A A C C T C C A G A T A G C T T C A G A A T 299
Qy 124 A s p a l a l e u l e s e r t r a r g r g l i l e u g l u a l a c y e g l u t h r l e u s e r a n p r o 143
Db 300 G A A G C T T G A T T A G T G T G T G C A G A T T C T T C A A G C T T T G T A A A C C T A G C T G A T C C T 359
Qy 144 A r g s e r a r g l u t y r a n g l u g l y l e u l e u a s p - - - - - A s p g l u g l u a l a 159
Db 360 G C T T C T A G A A G A G A T A T A T A C A A C C T C G T G A G A G A A C A A G A T G A G A A T C 419
Qy 160 T h r v a l l e t h r a s p a l p r o t r a s p l y s v a l p r o g l y a l a e u c y v a l l e u g l u 179
Db 420 T C C A T T C T A C T G A A A T C C C T T C G A C A A A G T T C T G A G A C T G T G C G T T G C A A G A A 479
Qy 180 G l y g l y l u t h r c l u l e v a l l e u a r g v a l g l y l u a l e u l e u l y s g l u a g l e u p r o 199
Db 480 G C T G A G A G C G A G T T G T G C T T G A T T G A G G G G T T T A C G A G A G A G A G A G T T A C C G 539
Qy 200 L y s e r p h e l y s e g l u a s p a l a l e u v a l l e u v a l m e t a l a l e u a l a p h e l e u a s p v a l s e r a r g 219
Db 540 A A G A T T T A G A C A A A T G T T G T G C T A T A G C C C T T G C A T A T T A C T T T C T T A G 559
Qy 220 A s p a l a m e t a l a l e u a s p p r o p r o a s p p h e i l e t h r g l y t r c l u p h e v a l g l u a l a 239
Db 600 G A T G C T A T G C C T T T G C C C G C A A T T C A T T G T T G T G A G A T G C T G A A A G G C A 659

```

Wed Feb 22 13:40:19 2006

us-10-600-070b-2.p2n.rnpbm

Page 18

Search completed: February 21, 2006, 17:04:36
Job time : 1617 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2006, 13:33:45 ; Search time 2594 Seconds
(without alignments)
655.503 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGTGLSPFQCRLLP.....YEVFWKSGMKITRGVTLAS 801

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7204323 seqs, 1061406715 residues
Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/abses/ABSSWEB.spool/US1060070/rnatc.21022006.112406.8805/app.query.fasta.1
-DB=Published Applications NA New -QFMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62
-TRANS-human0.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abses03p
-USER=US1060070 @CGN 1 1 335 @rnatc.21022006.112406.8805 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAITR -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:
1: /cgnt2_6/prodata/1/pubpna/US08 NEW PUB.seq:*
2: /cgnt2_6/prodata/1/pubpna/US07 NEW PUB.seq:*
3: /cgnt2_6/prodata/1/pubpna/US07 NEW PUB.seq:*
4: /cgnt2_6/prodata/1/pubpna/PCT NEW PUB.seq:*
5: /cgnt2_6/prodata/1/pubpna/US09 NEW PUB.seq:*
6: /cgnt2_6/prodata/1/pubpna/US09 NEW PUB.seq1:*
7: /cgnt2_6/prodata/1/pubpna/US10 NEW PUB.seq1:*
8: /cgnt2_6/prodata/1/pubpna/US10 NEW PUB.seq1:*
9: /cgnt2_6/prodata/1/pubpna/US11 NEW PUB.seq:*
10: /cgnt2_6/prodata/1/pubpna/US11 NEW PUB.seq2:*
11: /cgnt2_6/prodata/1/pubpna/US11 NEW PUB.seq3:*
12: /cgnt2_6/prodata/1/pubpna/US11 NEW PUB.seq4:*
13: /cgnt2_6/prodata/1/pubpna/US60 NEW PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133.5	3.3	5880	12	US-11-000-688-226
2	121.5	3.0	2454	8	US-10-793-626-49
3	121.5	3.0	2454	8	US-10-793-626-1527
4	121.5	3.0	3295	8	US-10-793-626-3676

Result No.	Score	Query Match	Length	ID	Description
5	121.5	3.0	3742	8	US-10-793-626-4076
6	121.5	3.0	3985	8	US-10-793-626-3856
7	120	3.0	12507	12	US-11-136-527-2447
8	116	2.9	4965	8	US-10-485-517-43
9	114.5	2.8	16317	12	US-11-117-187-212
10	114	2.8	3327	8	US-10-793-626-3414
11	112	2.8	1908	112	US-11-098-688-9088
12	112	2.8	1457619	12	US-11-098-688-8739
13	111.5	2.7	11115	8	US-10-513-786-6
14	111.5	2.7	11115	8	US-10-513-786-8
15	111	2.7	3252	12	US-11-136-527-3210
16	110.5	2.7	3389	9	US-11-072-512-842
17	110	2.7	3002	9	US-11-031-206-183
18	109.5	2.7	1983	8	US-10-467-657-1265
19	109	2.7	1209	8	US-10-858-730-58
20	109	2.7	6914	12	US-11-000-688-1053
21	109	2.7	10302	12	US-11-127-832-23
22	108	2.7	2724	8	US-10-467-657-1069
23	108	2.7	4767	12	US-11-000-688-952
24	108	2.7	4767	12	US-11-169-041-99
25	108	2.7	8503	12	US-11-124-368A-45
26	108	2.7	8512	12	US-11-124-368A-46
27	107	2.6	2358	12	US-11-055-822-225
28	107	2.6	2358	12	US-11-055-822-713
29	107	2.6	2358	12	US-11-055-822-755
30	107	2.6	2810	12	US-11-168-476-1
31	107	2.6	3158	9	US-11-031-206-187
32	107	2.6	10300	8	US-10-947-249-122
33	106.5	2.6	3060	8	US-10-793-626-3495
34	106.5	2.6	9286	12	US-11-136-527-578
35	106.5	2.6	10705	12	US-11-136-527-2232
36	105.5	2.6	3828	7	US-10-724-598-48
37	105.5	2.6	13144	7	US-10-724-598-41
38	105	2.6	3931	12	US-11-136-527-3988
39	105	2.6	11736	12	US-11-000-467-218
40	104.5	2.6	8730	12	US-11-087-100-1
41	104.5	2.6	8730	12	US-11-087-084-1
42	104.5	2.6	8730	12	US-11-087-085-1
43	104	2.6	1881	8	US-10-467-657-5431
44	104	2.6	2238	8	US-10-858-730-262
45	104	2.6	6012	8	US-10-467-657-83

ALIGNMENTS

RESULT 1
US-11-000-688-226
; Sequence 226, Application US/11000688
; Publication No. US20050287544A1
GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
SEQ ID NO 226
; LENGTH: 5880
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
NAME/KEY: misc_feature
; LOCATION: (1)..(5880)
; OTHER INFORMATION: eukaryotic translation initiation factor 4
US-11-000-688-226

Alignment Scores:

pred. No.:	0.00259	Length:	5880
Score:	133.50	Matches:	171
Percent Similarity:	33.3%	Conservative:	123
Best Local Similarity:	19.4%	Mismatches:	277
Query Match:	3.3%	Indels:	312
DB:	12	Gaps:	44

US-10-600-070b-2 (1-801) x US-11-000-688-226 (1-5880)

```

35 ThrilecySerAlaSerlySTrPAlaSPArg---LeuLeuSerAepPheAmPheThr 53
2102 ACTGATATCTAGAGTAAAGAGAGTATGACAGGAGTTTCTGCTGAGACTTCAGTTC--- 2158
54 SerAepSerSerSerSerSerPheAlaThrAlaThrThrAlaThrLeuValSerLeu 73
2159 -----ATG 2161
74 ProProSerIleAepArgProGluArgHleValProIleProIleAepPheTyrgLInVal 93
2162 CTGCTCTGATATACAAAGACAGAGGCTGCTCTCTATC----- 2200
94 LeuGlyAlaGlnThrHisPheLeuThrAepGlyIleArg-----Arg 107
2201 -----AGTATGTGGTTCTTGACAAATGCAACCAACCAATGCGCAATGCGA 2248
108 AlaPheGluAlaArgVal---SerLyPProProGlnPhe-----GlyPheSerAepAep 124
2249 ACTCTGATCTCTGAAATTTTGCTCGAGAGACAGACTTACACCAAGCTTTGCTGATTTT 2308
125 Ala-----LeuIleSerArgArgGln 131
2309 GGAAGCAGACACCTGCTGAGAGGCTTCTTGTGATGTGGGTGACCAAGA--- 2365
132 IleLeuGlnAlaIaLaCySGluThrLeuSerAepProArgSerArgArgGluTyrgAenGlu 151
2366 -----TCTCAACCTGGCCAAAGAGAAAGAACCAAGAG 2398
152 GlyLeuLeuAepAepGluGluAlaThrValIleThrAepValProTrpAepLyValPro 171
2399 ATCATC-----ACAGTTTCTGTAAAGAAAGATGTACCTGAAAAAGGCAGAA 2446
172 GlyAlaLeuCyValLeuGln-----GluGlyGlyGlu 182
2447 AATGCTTGGAAACCAAGCCAAAGAGAGACAGCAAGCCAGTATCCCGAAAACTTAA 2506
183 ThrGluIleValLeuArgValGlyGluAlaLeuLeuLySGluArgLeuProLySerPhe 202
2507 ACCAGAGAGCTTTTAAAGAAAGTTGAAAGTATCTTAATAATTAATTGACACCAAGATGTC 2566
203 LySGlnAepValValLeuValMetAlaLeuAlaPheLeuAepValSerArgAep----- 220
2567 AATCA-----CTGATGAGACAGAGTGTACAGACTTACTGTGACACAGAGACGG 2617
221 -----AlaMetAlaLeuAepProProAepPheIleThr 231
2618 CTGAAGAGAGTTATTGACCTGGCTTTGAGAAAGCTATGATGAACCACTTCTCTGTG 2677
232 GlyTyI----- 233
2678 GCTTACCAAAACATGTGTGCATGTCTAGTAAAGCTGAAGTAAAGTAAAGTAAAGTAA 2737
234 -----GluPheVal--- 236
2738 GGTAAACACAGTAAATTTCCGGAAGCTGTACTGAAACCGTTGCGCAAGAGAGTTGAAAA 2797
237 -----GluGluAlaLeuLyLeuLeuGlnGluGluGluAlaSerSerLeuAla 252
2798 GATTAAGCAGATGATGATGCTTTTGAAGAAAGAGAAAGAACTTGAAGGCTCCAGTGTCT 2857
253 ProAepLeuArgAlaGlnIleAepGluThrLeuGluGluIleThrProArgTyrgValLeu 272

```

```

2858 CCAGAGAGAGAGCAAGGCTTCAATGTAAGTGGAAAGAA----- 2896
273 GluLeuLeuGlyLeuProLeuGlyAepAepTyrgAlaAlaLyArgLeuAenGlyLeuSer 292
2897 -----GCCAAGAGCAAAAGCCGGCGGAGAA----- 2920
293 GlyValArgAenIleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGly 312
2921 -----TCCATTGGCAAC-----ATCAAGTTTATTGAGAA 2950
313 LeuThrArgGluLyPheMetAenGluAlaPheLeuAepMetThrAlaIaGluGlnVal 332
2951 CTTCTTAACTCAAAATGCTGACTGAAGCCATCATCAT----- 2989
333 AepLeuPheValAlaThrProSerAenIleProAlaGluSerPheGluValTyrgLInVal 352
2990 GACTGTGTGTGAAGCTGCTAAAGAACATGATGAAGAAATCCTGAGAGGCTGTGTGCG 3049
353 AlaLeuAlaLeuValAlaGln-----AlaPheIleGlyLyLeuPProHleLeuGln 370
3050 CTGCTCACCACCATGCGCAAAAGACTTGACCTTGAAGAAAGCAAGCCAGCTATG----- 3103
371 AepAlaAepLySGlnPheGlnGlnLeuGln-----GlnAlaLyValMetAlaMet 387
3104 -----GACCAGTACTTAAATCAGATGAGAGAAATTTGAAAGAAAGAAAGAAAGCTTCT 3157
388 GluIleProAlaMetLeuTyrgAepThrArgAenAenTrpGluIleAepPheGlyLeuGln 407
3158 AGGATTCGGTTCATCTCTCAAGATGTT-----ATAGAG 3190
408 ArgGlyLeuCyValLeuLeuIleGlyLyValAepGluCyArgMetTrpLeuGlyLeu 427
3191 CTAAAGCTGTGCAATTGGGATCTCGAAGAGCAGATCAA----- 3229
428 AepSerGluAepSerGlnTyrgAenProAlaIleValaGluPheValLeuGlnAenSer 447
3230 -----GGGCTTAAATCTATGACAGATTCACAAAGAGGCT 3265
448 AenAepAepAepAepAep----- 454
3266 AAAATTAAGAAACAAAGAAAGCAAGAGGCTCCAGCACTGACCAAGAAAGAGAGA 3325
455 LeuProGlyLeuCyValLeuLeuGln---ThrTrp-----LeuAlaGlyValAlaPhe 471
3326 AGACCAAGTGTCCAGAGAGTGAAGAGGAGGCTGGAACACTGACAAAGGGGCGCAAGAC 3385
472 ProArgPheArgAepThrLyAepLyValPheLyLeuGlyAepTyrgTyrgAepAepPro 491
3386 AGTGGGTACTGAC---CCCTCAAAATTCCTAAAAATCTAAGCCTTACATTTGAT--- 3439
492 MetValLeuSerTyrgLeuGluArgValaGluValaGlnGlySerProLeuAlaIaAla 511
3440 -----GAAAAATTCAGCTGTGACCTTAACCAACAGCTTACAGGACAGTGTG 3481
512 AlaThrMetAlaArgIleGlyAlaGluHleValIleValSerAlaMetGln----- 528
3482 GGAAGAGCAGAGCTGTGAGCAAGGCAAGTGAAGCTGATCCCTTACCGTCAAGTGTCT 3541
529 -----AlaLeuGlnLyValPheProSerArgTyrgTrpAepArg 541
3542 TCCAGTTTAAACAGATTCCTGCTGCAACCTCCAGACACCTCAGGCTCCAGCCATCC 3601
542 AenSerAlaGluProLyAepValaGlnGluThrValPheSerValAepProValaGlyAen 561
3602 ACCGCTGTAGAG---TTGATTTCCGAAGAGACTTAACTGATGTGTGAAGTATGAGGAGG 3658
562 AenValaGlyArgAepGlyGluProGlyValaPheIleAlaGluAlaValaPheProSerGlu 581
3659 GAG-----AAGATGACAGAGCC-----CTTCCATCTGCA 3688
582 AenPheGluThrAenAepTyrgAlaIleArgAlaGlyValaSerGluSerSerValAepGlu 601
3689 ACAGCTCGCCAAATATCTTTC---ATGAGGGGTGCGAGCATAAAGACTGCTGAGCAAT 3745

```

Query Match:	3.0%	Indels:	204
DB:	8	Gaps:	28
US-10-600-070B-2 (1-801) x US-10-793-626-49 (1-2454)			
QY 124	AspAlaLeuIleSerArgGlnIleLeuGlnAlaIaAcysGluThrLeuSerAnpro	143	
DB 538	GATCAGCGTCGACGACGAGATAAAGAAATTACTCGTAATTGAAGCTTTA-----	588	
QY 144	ArgSerArgArgGluTyrAsnGluGlyLeuLeuAspArgGluAlaThrValIleThr	163	
DB 589	---AGTCGCTGACTAATAAATATCTGCTGAATTTGGTGAA-----	627	
QY 164	AspValProTyrAspIysValProGlyAlaLeuCysValIleuGlnGluGlyGluThr	183	
DB 628	-----CCCGGCGTT-----	645	
QY 184	GluIleValLeuArgValGlyGluAlaLeuLeuGlyArgLeuProLysSerPheLys	203	
DB 646	GCAATTCGTGAAGGCGCTTCCGACACCATTTGTTAAATAATGAAGTCCGAAACTTTAAA	705	
QY 204	GlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAla	223	
DB 706	-----GACAAAGCTAATGTCA-----	723	
QY 224	LeuAspProProAspPheIleThrGlyTyrLeuPhe-----ValGluGluAlaLeu	240	
DB 724	TTAAGATATGGGTACAGCTCGTAGCTGGACCTAAATATCGTGGAATTTGAAAGAAATTGG	783	
QY 241	LysLeuLeuGlnGlnGlu-----	246	
DB 784	AAAAAAGTTATGAGAGAAATCATCAACGCTGGTATTTATTCTATTATTCATGATCACT	843	
QY 247	-----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspIleThr	262	
DB 844	CATACCTTTAGTTGGGCGCTGGTGGCGACAGAGAGCAATTGATGCA-----TCTAATATT	897	
QY 263	LeuGlnGluIleThrProArgTyrValIleuGluLeuLeuGlyLeuProLeuGlyAspAsp	282	
DB 898	TTAAAAACCTGCTTAGCTCGTGAGAAATGCAATGATATATAGGTCACAACTTAGATGAA	957	
QY 283	Tyr-----AlaIleLysArg-----	287	
DB 958	TATGCTAAAAATATAGAAAAAGACGCTGATTCAGACCTCGTTTCAACCAATTCACGTG	1011	
QY 288	-----LeuAsnGlyLeuSerGlyValAlaArgAsn	296	
DB 1018	GATGAACCTACAGTTGAAGACACAGCATTTGAATCTTTAAAGATTTACGACCGTTATAG	1077	
QY 297	IleLeuTyrSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlu	316	
DB 1078	GCTCATCACAGAAATTAATATCTCAGATGAGACCTTTAGAGCGGCGTCTAAATTGAGTGAT	1133	
QY 317	LysPheMetAsnGluAlaPheLeuArgMetThrIleAlaGluGlnValAspLeuPheVal	336	
DB 1138	CGCTATGTTTCAGATCGTTCTTCGACAGATAAACCCATTCCTTAATGATGAGGCAAGT	1197	
QY 337	Ala-----ThrProSerAsnIleProAlaGluSerPheGlu	348	
DB 1198	TCAAAAGTACCTTAAAAAGTCATACACGCGCAAGTATTTAAAAAGAGATTGACACAGAA	1255	
QY 349	ValTyrGluValAla-----LeuAlaLeuValAlaGlnAlaPheIleGlyLys	364	
DB 1258	ATTGATTAAGTAAATAATGAAAAAGATGCTCAGTTCATGCTCAAGAATTT-----	1308	
QY 365	LysProHISleuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnGlnAlaLysVal	384	
DB 1309	---GAAATGCGCGCTAAATTTAAGATAGCAATCTTAACCTTGAAAGCAATATGAA---	1363	
QY 385	MetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTyrGlu-----	401	
DB 1363	-----GATCTAATAAATGATGAGAAAAATGCACA	1392	

```

Qy 402 -----IleAspPheGlyLeu---GluArgGlyLeuCysAlaLeuLeuIleGlyLysVal 418
Db 1393 GGTGGTTAGATACGCTTATCTGAAAGAAATATCCCTGAAGTAAGTGGT----- 1446
Qy 419 AspGluCysArgMetTrpLeuGlyLeu-----AspSerGluAspSer 432
Db 1447 -----TCGACAGGATATCTTTAACTAAATTAATGAACTGAATCAGAT 1491
Qy 433 GlnTyrArgAsnProAlaIleValGluPheValLeuGluAsnSerAsnAgaAspAsn 452
Db 1492 CGTTTATTGAAT-----CTTGAAGATACACTTAACCTGCTCATTCGACAAAC 1542
Qy 453 AspAspLeuProGlyLeuCysLysLeuLeuGluThrTrpLeuAlaGlyValIlePhePro 472
Db 1543 GATGCTGTCAATCAATTAATGAAGCTTGAAGAGCTGCTGCTCTT----- 1593
Qy 473 ArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr---TyrAspAspPro 491
Db 1594 -----AAAGATCCAAACGTCCAATCGTAGTTTATTTCTTAGGACCT 1638
Qy 492 MetValLeuSerTyrLeuGluArgValGluValIleGlnIleSerProLeuAlaIleAla 511
Db 1639 ACAAGTGTGGTAAACCTGAATGGCTGCTGCTTACCTGAATCTATGTTGCTGAAGAC 1698
Qy 512 AlaThrMetAlaArgIleGly---AlaGluHileValLysAlaSerAlaMetGlnAlaLeu 530
Db 1699 GATGCAATGATTCGCGTGAATGATGAGTAATTAATGAGAAACATGCTGACGTGCTTA 1758
Qy 531 GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
Db 1759 GTTGTCACCTCCAGATATGTAGACATGATGACGCGCTCAATGATGAAAGTT 1818
Qy 543 SerAlaGluProLysAspVal-----GlnGluThrVal 553
Db 1819 AGACGTAAACCATATCTGTGATTTATTTGATGAATTAAGAAACATCTCGACGTA 1878
Qy 554 PheSerVal-----AspProValGlyAsnAsn 562
Db 1879 TTTAATTTCTTCTACAGTTTATGATGATGATTAACAGTACTTAAGTGGTACT 1938
Qy 563 ValGlyArgAspGlyLysProGlyValPheIleAlaGluAlaValaLysProSerGluAsn 582
Db 1939 GTG-----GACTTCGCTAATACTGATGATTAATGACTTCTAATGCGGACCAAGAA 1992
Qy 583 PheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerValAspGluThr 602
Db 1993 TTAACAGACCAACGCTTGTGCTGTTTGAAGTCTTCAGAAAGGTAGT---GACTACGAA 2049
Qy 603 ThrValGluMetSerValAlaAspMetLeuLysGluAla----- 615
Db 2050 ACTGTTCAGAAAACAAATGATGAAGAAATTAATAATTCATTCGACCAAGATTTCTTAAC 2109
Qy 616 -----SerValLysIleLeuAla 621
Db 2110 CGTGTGATGATATTATGCTTCTCCAAACCTTACAAAGATGATTAAGAAATTTGTT 2169
Qy 622 AlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLys-----Tyr 636
Db 2170 ACAATGATGCTAAATTAACCTTACCTGCTTTCAGAGCAAAATATTAATATTGTTGTT 2229
Qy 636 rPheLeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
Db 2230 ACTGATTAAGCGAAGAAAGAAATTCGACAGACGATATGATCTGATATATGTCCTAGA 2289
Qy 656 p 656
Db 2290 C 2290

```

```

RESULT 3
US-10-793-626-1527
; Sequence 1527, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:

```

```

; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1527
; LENGTH: 2454
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1527

Alignment Scores:
Pred. No.: 0.0177 Length: 2454
Score: 121.50 Matches: 125
Percent Similarity: 33.0% Conservative: 93
Best Local Similarity: 18.9% Mismatches: 239
Query Match: 3.0% Indels: 204
DB: 8 Gaps: 28

US-10-600-070b-2 (1-801) x US-10-793-626-1527 (1-2454)

Qy 124 AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCysGluThrLeuSerAsnPro 143
Db 538 GATCCAGTGTGATGACGAGATTAAGAAATTAATCTCGTAAATGTAAGCTTTA----- 588
Qy 144 ArgSerArgArgGlnTyrAsnGlnGlyLeuLeuAspGlnLysValaThrValIleThr 163
Db 589 ---AGTCGTCGTACTTAATAATATCTGCTAATTTGGTGA----- 627
Qy 164 AspValProTrpAspLysValProGlyAlaLeuCysValLeuGlnGluGlyGluThr 183
Db 628 -----CCCGGTGTT-----GGTAAACA 645
Qy 184 GlnIleValLeuArgValGlyGluAlaLeuLeuGlyLysArgLeuProLysSerPheLys 203
Db 646 GCAATGTGCGAAGGCTGTGGCAGACAAATTTGTAAGAAATGACACGAACTTTAAAA 705
Qy 204 GlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAla 223
Db 706 -----GACAAAGCTGATATGTC 723
Qy 224 LeuAspProAspPheIleThrGlyTyrGluPhe-----ValGluGluAlaLeu 240
Db 724 TTAGATATGGGTACAGTCGATGCTGACACTTAATATCTGTGATTAATGAAAGATTTG 783
Qy 241 LysLeuLeuGlnGluLys----- 246
Db 784 AAAAAAGTTATGAGAGAAATCCATCAAGCTGTAAATGTTATTTATTCATGAACTT 843
Qy 247 -----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThr 262
Db 844 CATACTTTAGTTAGTGGCCGTGTGGCCGACGAAGACGAATGATGCA-----TCTAATATT 897
Qy 263 LeuGlnGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAsp 282
Db 898 TTAACACCTGCTTACGTCGTGAGAAATTCATGATATGATGTCACCAACATTAGATGAA 957
Qy 283 Tyr-----AlaAlaLysArg----- 287
Db 958 TATCGTAAATAATATGAAAAAGACGCTGCTTAAGACGCTGTTTCAACCAATTCAAGTG 1017
Qy 288 -----LeuAsnGlyLeuSerGlyValArgAsn 296
Db 1018 GATGAACCTACAGTTGAAGACAGATTTGAATCTTTAAAGATTAAGTACCGTTATGAG 1077
Qy 297 IleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyLysLeuThrArgGlu 316

```



```

Db      2743 TTAGATATGGGTACAGTCTGACGTGACACTAAATATCTGTGTGTAATTTGAAAGAAATG 2684
      241 LysLeuLeuGlnGlnLys----- 246
      2683 AAAAAAGTTATGAGAAATCCATCAAGCTGTAAATGTTATCTTATTTATCGATGAACCT 2624
      247 -----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGlnThr 262
      2623 CATACTTACTAGTGGCCGTGTGGCGCAGAAAGCAATTGATGCA-----TCTAATATT 2570
      263 LeuGlnGlnIleThrProArgTyrValLeuGlnLeuGlnGlyLeuProLeuGlnAspAsp 282
      2569 TTTAAACCTGCTTAGCTGTGAGAAATTCATGTATAGTGTCCCAACATTAGATGAA 2510
      283 Tyr-----AlaAlaLysArg----- 287
      2509 TATCGTAAATAATATAGAAAAAGACGCTGATTAAGACGTGTTTCAACCAATTCAAGT 2450
      288 -----LeuAsnGlyLeuSerGlyValArgAsn 236
      2449 GATGAACCTACAGTTGAGACAGATTTGAATCTTAAAGATTAACGTGACCGTTATGAG 2390
      297 IleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGln 316
      2389 GCTCATCAACAGATTATATCTCAGATGAGCTTTAGAACGCGCTGCTAATTTGAGTAT 2330
      317 LysPheMetAsnGlnAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheVal 336
      2329 CGCTATGTTTCAGATCGTTCCTTGCCAGATTAAGCCATTGACCTTAATGATGAGCGCA 2270
      337 Ala-----ThrProSerAsnIleProAlaGlnIleAspLeuPheGln 348
      2269 TCAGAACTTAGACTTAAAGATCATACACCGCAAGTAATTTAAAGAGATTTGAACAGAA 2210
      349 ValTyrGlnValAla-----LeuAlaLeuValAlaGlnAlaPheIleGlyLys 364
      2209 ATGTATTAAGCTAAATAAAGATAAAGATGCTGAGTCTCATGAGTAATTT----- 2159
      365 LysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnGlnGlnAlaLysVal 384
      2158 ---GAAATATCCGCTAATTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2105
      385 MetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAspTrpGln----- 401
      2104 -----GATCTAAATAATGATGAGAAATAATGACAA 2075
      402 -----IleAspPheGlyLeu---GluArgGlyLeuCysAlaLeuLeuIleGlyLysVal 418
      2074 GGTGTTTATGATATCTGCTTATCTGAAAGAAATATCTGCTGAAATTAAGCTGCTG 2021
      419 AspGlnCysArgMetTrpLeuGlyLeu-----AspSerGlnAspSer 432
      2020 -----TGGACAGGTATTCCTTTAACTAAATTAATTAAGAACTGAATTCAGAT 1976
      433 GlnTyrArgAsnProAlaIleValGlnPheValLeuGlnLeuSerAsnArgAspAsn 452
      1975 CGTTTATTTGAAT-----CTTGAAATTAACACTTCAATAACGTGATTTGACAAAC 1925
      453 AspAspLeuProGlyLeuCysLysLeuLeuGlnThrTrpLeuAlaGlyValValPhePro 472
      1924 GATGCTGTCAATTAATTAAGAACTGTGAGAAAGCTGCTGCTGCTCTT----- 1874
      473 ArgPheArgAspThrLysAspLysLysPheLysLeuGlnAspTyr---TyrAspAspPro 491
      1873 -----AAAGATCCAAACGCTCAATCGGTAGTTTATTTTCTTGAACCT 1829
      492 MetValLeuSerTyrLeuGlnLysValGlnValGlnGlySerProLeuAlaAlaAla 511
      1828 ACAGGTGTGGTAAACGAAATGCTGCTGCTTACGTGATATCTAATTTGTTGGAAGAC 1769
      512 AlaThrMetAlaArgIleGly---AlaGlnHisValLysAlaSerAlaMetGlnAlaLeu 530
      1768 GATGCAATGATTCGCGTATGATATGATGAATTTATGAGAAACATGCTGCTGATGATTA 1709

```

```

Qy      531 GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
      1708 GTTGTGACACCTCCAGATATATGATGAGACATGATGACGCGGTCAATTGACTGAAAAAGTT 1649
      543 SerAlaGlnProLysAspVal-----GlnGlnThrVal 553
      1648 AGACGTAAACCAATCTCTGTGATTTTATTTGATGAATAATGAGAACCATCTCTGACGTA 1589
      554 PheSerVal-----AspProValGlyAsnAsn 562
      1588 TTTAATATTTCTTTACAGTTTAAATGATGATGATGATTAACAGATTAAGTCTGATCT 1529
      563 ValGlyArgAspGlyLysProGlyValPheIleAlaGlnAlaValArgProSerGlnAsn 582
      1528 GTG-----GACTTCGCTAATACGTGATATATGATGATCTTAATGTTGGAGCTCAAGAA 1475
      583 PheGlnThrAsnAspTyrAlaIleArgAlaGlyValSerGlnSerSerValAspGlnThr 602
      1474 TTACAGACCAACGCTTGTGCTGTTTGGAGGTGCTTCAGAGTGT---GACTACGAA 1418
      603 ThrValGlnMetSerValAlaAspMetLeuLysGlnAla----- 615
      1417 ACTGTCAAAAAACATGATGAAAGATTTAAATAATTCATCCGACCAAGATTTTAAAC 1358
      616 -----SerValLysIleLeuAla 621
      1357 CGTGTGATGACATTAATGCTTCCACAACTTAACAAAGATGAATTAAGAAATTTGTT 1298
      622 AlaGlyValAla-IleGlyLeuIleSerLeuPheSerGlnLys-----Tyr 636
      1297 ACAATGATGTAAATTAATCTACTACCGCTTTCAGAGCAAAATTAATTAATTTGTTGTT 1238
      636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAs 656
      1237 ACTGATTAACCGAAAGAAAAAATTGCAAGAAAGATATGATCTGAATTTGTCGTAGA 1178
      656 P 656
      1177 C 1177

```

```

RESULT 5
US-10-793-626-4076/c
; Sequence 4076, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OR INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4076
; LENGTH: 3742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4076

```

```

Alignment Scores:
Pred. No.: 0.0298 Length: 3742
Score: 121.50 Matches: 125
Percent Similarity: 33.0% Conservative: 93
Best Local Similarity: 18.9% Mismatches: 239
Query Match: 3.0% Indels: 204
DB: 8 Gaps: 28

```

US-10-600-070B-2 (1-801) x US-10-793-626-4076 (1-3742)


```

Db      962 CGTGTGATGACATTATGTTCTTCACAACTTACAAAGATGATTAAAAAATTGTT 903
Qy      622 AlAGlyValAla-1leGlyLeu1leuSerLeuPheSerGlnlys-----Ty 636
Db      902 ACAATGATGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTGTT 843
Qy      636 rPheLeu1leuSerSerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAs 656
Db      842 ACATGATTAAGCGAAAGAAAATTTGCGAAGAGAGATATGATCTCTGATATGTCGTCAG 783
Qy      656 p 656
Db      782 C 782

RESULT 7
US-11-136-527-2447
; Sequence 2447, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2447
; LENGTH: 12507
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (12245)..(12245)
; OTHER INFORMATION: n i s a, c, g, or t
US-11-136-527-2447

Alignment Scores:
Pred. No.: 0.193 Length: 12507
Score: 120.00 Matches: 146
Percent Similarity: 32.7% Conservative: 98
Best Local Similarity: 19.5% Mismatches: 313
Query Match: 3.0% Indels: 190
Db: 12 Gaps: 29

US-10-600-070b-2 (1-801) x US-11-136-527-2447 (1-12507)
Qy      53 ThisSerAspSerSerSerSerSerPheAlaThrAlaThrThrAlaThrLeuValSer 72
Db      1986 ACTCTTAAGCAAAAGCGGGGTGAAGAGAGACCGACCTGCCACCCAGTCCTAAGCCT 2045
Qy      73 LeuProProSerLeuLeuArgProGluArgHisValProLeuProLeuLeuArgPheTyGln 92
Db      2046 GTTCCAGAGCTCCCAAGAGCGGGAGGCTGAGGAACCTGCTCCCAAGCCTTACTCTCAG 2105
Qy      93 ValLeuGlyAlaGlnThrHisPheLeuThrAspGlyLeuArgAlaPheGluAlaArg 112
Db      2106 GATCTGTCTGAAGCCACCAAGCCTCAGTAT----- 2138
Qy      113 ValSerLeuProProGlnPheGlyPheSerAspAspAlaLeuLeuSerArgArgGln--- 131
Db      2139 -----ACAGGCTACTCTCTCCGATGCGCTCTCCAGCTCCAGAGTGA 2180
Qy      132 -----1leuGlnAlaAlaCyGluThrLeuSer----- 141
Db      2181 ATCAACAGTCTTTATACGACGAAGTGAACACTTACACAGTCCGGGGGTGACAGGCCCA 2240
Qy      142 -----AsnProArgSerArgArgGluTyArgGlnGlyLeuLeuAspArgGln 158
Db      2241 CGTCAACCCAGCCCTCAGAGCTCCACAAAGTGGGAGTAGTATGCGCCTTCCCTGAG 2300

```

```

Qy      159 Ala---ThrVal1leuThrAspValProTrpAspLysValProGlyAlaLeuCyValLeu 177
Db      2301 GCCAGGCTGTGCTCCAGTGTGAGTGAAGAGACCAAGCCAGCGGAGCTGTGA 2360
Qy      178 GlnGlnGlyGlyGlnThrGlu1leuValLeuArgValGlyGlnAlaLeu---LeuArgLys 196
Db      2361 GACCAGAAAGCGGCTCCCACTCCCTCCATCATGACCCAGGCTTCGACTGACAG 2420
Qy      197 ArgLeuProLysSerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAsp 216
Db      2421 GAGCTGGGGATATTCTTAAGAGAGTACTCTCTGGCTATGGG-----CGTCAG 2471
Qy      217 ValSerArgAspAlaMetAlaLeuAspProProAspPheLeuThr-----GlyTyGln 234
Db      2472 AGGAGACGACGAACACACAGCTGACTTTCAGATGACTTTGGCAGCCAGCTGAGCAGAC 2531
Qy      235 PheValGlnGlnAlaLeuLysLeuLeuGlnGlnGlnGlyAlaSerSerLeuAlaProAsp 254
Db      2532 TACGTGGAGACAGC-----AGTGAAGCGGCTGTGCTCCCTCCACCTCAG 2579
Qy      255 LeuArgAlaGln1leuAspGluThrLeuGlnGln1leuThrProArgTyValLeuGlnLeu 274
Db      2580 CCCCAGCCAGGCGAGATGACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2639
Qy      275 LeuGlyLeuPro-----LeuGlyAspAspTyAlaAlaLysArgLeuAsnGlyLeuSer 292
Db      2640 AGCGCCGAGGAACAACTTGGAGAGAGACACCGCTTCTCCGGCCAGCCCTGACC 2659
Qy      293 GlyValArgAsn1leuLeuTrpSerValGlyGlyGlyValAsnAlaLeuValGlyGly 312
Db      2700 -----AAACACGGTGCCAGAAAGCAAGTGCC----- 2726
Qy      313 LeuThrArgGlnLysPheMetLeuGlnAlaPheLeuArgMetThrAlaAlaGlnGlnVal 332
Db      2727 AGACCAAGGCCCAATCTAGCCAAAGATCA----- 2756
Qy      333 AspLeuPheValAlaThrProSerAsn1leuProAlaGlnSerPheGluValTyGlnVal 352
Db      2757 -----GTAGCACTGCCCAAGAG----- 2774
Qy      353 AlaLeuAlaLeuValAlaGlnAlaPhe1leuGlyLysLeuProHisLeuLeuGlnAla 372
Db      2775 -----CGCTGCCCAT-----AATGCC 2792
Qy      373 AspLysGlnPheGlnGlnGlnGlnGlnAlaLysValMetAlaMetGluLeuProAlaMet 392
Db      2793 ACCACAGTTACGAGAACTT-----CTGCTGAGAAAGGCCAGCAGAG 2837
Qy      393 LeuTyAspThrArgAsnAsnTrpGlu1leuAspPheGlyLeuGlnArgGlyLeuCyAla 412
Db      2838 -----CCCACTGACGAGCCCTTCAAGGAGGATCTCCAGCC 2873
Qy      413 Leu---Leu1leuGlyLysValAspGluCyArgMetTrp-----LeuGly 426
Db      2874 TTTAAACATTTATGATCAACACACACAGATGATGATGATGATGATGATGATGATGATGAT 2933
Qy      427 LeuAspSerGlnAspSerGlnTyArgAsnProAla1leuValGluPheValLeuGluAsn 446
Db      2934 CAAGGCTCTGATCCCAACCTAGACCGGAGCCTGAGCTGAGATGAGAGACCTTACAGCC 2993
Qy      447 SerAsnArgAsp----- 450
Db      2994 TCGCTGAAAGACCGCTCTCTGAGGAGAGACTCTTACACTGCGCGCTTCCAGCCTTACG 3053
Qy      451 -----AspAsnAspLeu 455
Db      3054 TATACATCGGACACTCACCACTCCCTGCTTCCCTGAGAGAGACAGTGA CAGAGC 3113
Qy      456 ProGlyLeuCyLysLeuLeuGlnThrTrpLeuAla----- 467
Db      3114 CCGAGCGGAGGAGGCTTAGAAAGAGCAAGCAGCGGAGCCCGGACCGGATCC 3173
Qy      468 ---GlyValValPheProArgPheArgAspThrLysAspLysLysPheLysLeuGlyAsp 486

```



```

; LENGTH: 3327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3414

Alignment Scores:
Pred. No.: 0.169 Length: 3327
Score: 114.00 Matches: 126
Percent Similarity: 33.8% Conservative: 98
Best Local Similarity: 19.0% Mismatches: 214
Query Match: 2.8% Indels: 224
DB: Gaps: 31

US-10-600-070b-2 (1-801) x US-10-793-626-3414 (1-3327)
QY 71 ValSerLeuProPseSerIleAspArgProGluArgHisValProIleProIleAspPhe 90
DB 869 TTAACCTTAATCCCTCAGTTGAT-----ATATCTTATCCTTTAGAT--- 910
QY 91 TyrGlnValLeuGlnAlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGlu 110
DB 911 -----CACTTAATTAGATACCTGTTAATAGGGA----- 940
QY 111 AlaArgValSerIleProProGlnPheGlyPheSerAspAspAlaLeuIleSerArgArg 130
DB 941 TCTCAACACAGCAAA----- 955
QY 131 GlnIleLeuGlnIleAlaCysGluThrLeuSerAsnProArgSerArgGluTyrAsn 150
DB 956 -----ACAGCAGCGCGTAAAGCATTAATGTTACTAGATTTGCTGAATTTGCT 1006
QY 151 GluGlyLeu----- 157
DB 1007 GAAGACGTATAGCAAGTGTGTTTGTAGTGAGCATTAGTCATATATTGAAGAACAA 1066
QY 158 GluAlaThrValIleThrAspValProTrpAspIleValProGly-----Ala 173
DB 1067 ATGAAACATACCTGATATTMAACAAGCATTTTTCAMAAATCAAGCGCAACAGCAATTC 1126
QY 174 LeuCysValLeuGlnGlnGlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeu 193
DB 1127 ATTCGATACCTACCTAGAGCAACAAGCACTGAATCTTGAAGAGGCTCCACAGATTGAA 1186
QY 194 LeuIleGlu-----ArgLeuProIleSerPheIleGlnAspValVal 207
DB 1187 CTTAAGGAATCAGAGCAATTCAGTCACATCTATTAAACCTTTCAAGAAAGCATGATG 1246
QY 208 LeuValIleAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProPro 227
DB 1247 GCTGTATGCTGCTGTAGTCTG-----CCCAAA 1273
QY 228 AspPheIleThrGlyTyrGluPheValGluGluAlaLeuIleLeuGlnGlnGly 247
DB 1274 GGACTTAATACCTATTAAT-----TATCGGATATGTTGATTTAGCAAAAGCAAGCA 1327
QY 248 AlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeu-----GluGluIle 266
DB 1328 ATTTGACCATTTTAGTAGCTGCTGCTGATCATCACTTGAGAGAGTCTTATTAGTATG 1387
QY 267 ThrProArgTyrVal----- 271
DB 1388 AAACCTACAGTATTTAAACCAATATATAGATTAATCATCACTTTTAATTACAAAGTA 1447
QY 272 -----LeuGlu 273
DB 1448 ACCAATGATATTAAAGATTGAAGCGCGAGTAAGTCACCAATATTTAATGATTTGAA 1507
QY 274 LeuLeuGlyLeuProLeuGlyAspAspTyrAlaAlaIleArgLeuAsn----- 289
DB 1508 TGATATTATGTTTCATTGGCGAGTGAAGGCTTTTTCAGAAACATATATCAAAATTTTAT 1567
```

```

QY 290 -----GlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGly----- 305
DB 1568 AAGGTGAATATTCCTCCCAACATTAA-----GTAGTTAATCTGTGTGGTCAGAGATTCCACT 1624
QY 306 -----AlaSerAlaLeuValGlyGlyLeuThrArgGlyAspPheMetAsnGlu 321
DB 1625 GTAGCAGCAATTCCTTGTGACCTCATTCATTCACCAACAAACGATGAGATTAATAAAAA 1684
QY 322 -----AlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrPro 339
DB 1685 GCAAATGCATTCGGAATGCTAAATGCAATGCAACAA-----CAACA 1726
QY 340 SerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuAlaGln 359
DB 1727 GGTCAATTTAATACAGATTAATTTAGCAAAATATTACAAACAAATAGAAAGTTATAGAGCTG 1786
QY 360 AlaPheIleGlyIleIleAspProHisLeuLeuGlnAspAlaAspIleGlnGlnLeu 379
DB 1787 TAATTTATGCAAAA----- 1801
QY 380 GlnGlnAlaIleValIleMetAlaMetGlu-----IleProAla 391
DB 1802 TCACAAACAAAGAGTGCATCAATTAAGAAATTAAGTATCAGAAAGGTATTATTACAGCT 1861
QY 392 MetLeuTyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGluArgGlyLeuCys 411
DB 1862 TTAGCATTTGAT-----CAACGTGTGCATTA 1888
QY 412 AlaLeuLeuIleGlyIle-----ValAspGluCysArgMetTrp 424
DB 1889 AAAAGATGATGTCAGCAACATCAATCTGAAACACCAACAGTTGAACAAATAGAACATTA 1948
QY 425 LeuGlyLeuAspSerGluAsp-----SerGlnTyrArgAsnProAlaIleVal-----Glu 441
DB 1949 AAGTACTGTTCTTCTAGAGATTAATCAATATGCGCTTCATATTATTATAGATCCAGAA 2008
QY 442 PheValLeuGluAsnSerAsnArgAspAspAsp----- 453
DB 2009 TATGTTTACAGCAACATCAGATGCTCGAATTAAGTACGCGGACTATTATCTGCATACGAA 2068
QY 454 -----AspLeuProGlyLeuCysIleLeu-----LeuGluThrTrpLeu 466
DB 2069 AAAACGTATGATGATGATGGAAGGCTGTTGCCAGATTGCTTGTGATGATGCT 2128
QY 467 AlaGlyValAlaPheProArgPheArgAspThrIleAspIleIlePheIleGlyAsp 486
DB 2129 GCG-----AAACGTTGAAGAGCAAGGCGCCAAATGCAGTTAAATTTTAACTT 2176
QY 487 TyrTyrAspAspProMetValLeuSerTyrLeuGluArgValGluValValGlnGlySer 506
DB 2177 TATTATGAT-----GTAGATGACACAGAAGAATT 2206
QY 507 ProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGluIleValIleValIleSerAla 526
DB 2207 AACATCAAAAGCAACATATATTTGAACGAAATTTGTTGCAAAATGTTGCCGAAGATATT 2266
QY 527 MetGlnAlaLeuGlnIleVal-----PheProSerArgTyrThrAspArgAsnSerAlaGlu 545
DB 2267 CTTTCTCTTGGAAGTTTAAATATATGACGACAAATATTCCTGACATTAAGTGCAGAA 2326
QY 546 -----ProIleAspValGlnGluThrVal----- 553
DB 2327 TTCGCTAAAGTTAAGCCAGATTAAGTTAATGAAGCAATGAATTAATCTCTGAAGATCGT 2386
QY 554 PheSerValAsp-----ProValGlyAsnAsnValGlyArgAspGly 567
DB 2387 TTTAATGTGATGATCTTAAGTTGAAGTACCTGTGAATATATGAAATTTTGTGAAAGATTT 2446
QY 568 GluProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
DB 2447 TCAGAGGAGGAAGTGTATTACTTAAGAAAGAGTGCACAAACATTTCCGTGATCAAGAT 2506
```



```

Db      1591 -----TGGGCACATTCTAGAACTTTAGTGAAGTACGGCTCGATTGTTGATAGT 1641
               |||
               |||
Qy      732 ThrlenuleuylserValaspSer-----ValThrValSer 744
               |||
               |||
Db      1642 GAAGTAAAAAATTAAGAAAGCTAAGAAAAATGTCAGAACTATTCACAAAATAT 1701
               |||
               |||
Qy      745 AlaapglYThrArGAlaValGluAlaThrleuGluGluSerAlaCySleuSerAsp 764
               |||
               |||
Db      1702 CTTGATACGTTGCTGCTCTTCTGCTACAGCACTTGAACGAAACATTACAGGTGAT 1761
               |||
               |||
Qy      765 LeuValHis-----ProGluAspAsnAla 772
               |||
               |||
Db      1762 GATATATATCTTCTTATCAAGGGGAGAACTACCTCCATTATATACGAAACTCTAAT 1821
               |||
               |||
Qy      773 ThrAspValArGThrYrThrThrArGTYrGluValPhe 785
               |||
               |||
Db      1822 ACTGAAATTAATAAAGAAAGATTCAAATATCTGAAACCTTT 1860
               |||
               |||
RESULT 12
US-11-098-686-8739
; Sequence 8739, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OR INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098, 686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8739
; LENGTH: 1457619
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8739
Alignment Scores:
Pred. No.: 515 Length: 1457619
Score: 112.00 Matches: 124
Percent Similarity: 34.3% Conservative: 86
Best Local Similarity: 20.2% Mismatches: 213
Query Match: 2.8% Indels: 190
DB: Gaps: 32
US-10-600-070b-2 (1-801) x US-11-098-686-8739 (1-1457619)
Qy      235 PheValGluGluAlaLeuLeuLeuGluGluGluGluAlaSerSerLeuAlaProAsp 254
               |||
               |||
Db      237374 TTTGGTCGCTCAGGTGCAAGAAATGCTTAACACAGAAACAAGAGATCACATTGAAGAT 237433
               |||
               |||
Qy      255 LeuArGAlaGluAlaLeuSerGluThrLeuGluGluThr----- 267
               |||
               |||
Db      237434 GTT---GCAAGTGTGATGTAAGCTTAAGAACTCTCAAGTTGATTTTTTATCT 237490
               |||
               |||
Qy      268 -----ProArGTYrValLeuGluLeuLeu 275
               |||
               |||
Db      237491 AATCCAGAAAAATTTACACGATTAGTGTGCGATACCTTAAGAGATCTT---CTTGTT 237547
               |||
               |||
Qy      276 GlyLeuPro---LeuGlyAspAspTYrAlaAlaLysArGLeuAsnGlyLeuSerGlyVal 294
               |||
               |||
Db      237548 GGCCTCCTGCTGACAGGAAAAACATTCCTGCTCAGCTGTAGCTGTGTAAGCAGGGGTT 237607
               |||
               |||
Qy      295 ArgAsnIleuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThr 314
               |||
               |||
Db      237608 CCT-----TTTTCATTCATTTCTGCTTGAT----- 237634
               |||
               |||
Qy      315 ArgGluLysPheMetArgGluAlaPheLeuArgMetThrAlaGluGluValAspLeu 334
               |||
               |||

```

```

Db      237635 -----TTTGTGAGATGTTTGTGTGTGAGACATCTACAGATCTC 237682
               |||
               |||
Qy      335 PheValAlaThrProSerAsnIleProAlaGluSerPheGluValYrGluVal--Ala 353
               |||
               |||
Db      237683 TTTATACAGGTAAAAAATGCTCCATGCTTATTTT---ATTGATGAATTTGAATGCA 237739
               |||
               |||
Qy      354 LeuAlaLeuValAlaGluAlaPheIleGlyLysLysProHisLeuLeuGluAlaAsp 373
               |||
               |||
Db      237740 GTTAGACGTACGCTGCGCTGCGCTTGCC-----GAGACATGAT 237781
               |||
               |||
Qy      374 LysGluPheGluGluGluGluGluAlaLysValMetAlaMetGluIleProAlaMetLeu 393
               |||
               |||
Db      237782 GAAGAAGACAAACACTTATCACTTTAGTT----- 237814
               |||
               |||
Qy      394 TyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGluArgGlyLeuCySAlaLeu 413
               |||
               |||
Db      237815 -----GAAATGAT---GGTTTGAATCTAAGAAAGGTGTTAT 237850
               |||
               |||
Qy      414 LeuIleGlyLysValAspGluCySArgMetTrpLeuGlyLeuAspSerGluAspSerGln 433
               |||
               |||
Db      237851 TTAATCGACACTCAAAAT----- 237868
               |||
               |||
Qy      434 TyrArgAsnProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAsnAsp 453
               |||
               |||
Db      237869 -----CGTCAATGTTCTTATTCAGCTCTTAAGACCTGGAAGCTTTGATAGCAA 237922
               |||
               |||
Qy      454 -----AspLeuProGlyLeuCySLeuLeuGluGluThrTrpLeuAla 467
               |||
               |||
Db      237923 GTTGTGTTCTTCTTCTCCTGATGTAAAGACGTTTAAATCTTGAA----- 237970
               |||
               |||
Qy      468 GlyValAlaPheProArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTYr 487
               |||
               |||
Db      237971 -----GTCCATCTGCT---CGAACCACTTGATTAAGATGTTATCTT----- 238012
               |||
               |||
Qy      488 TyrAspAspProMetValLeuSerTYrLeuGluArgValGluValAla---GinglySer 506
               |||
               |||
Db      238013 -----GAAATTATAGCTAAGACGACA 238033
               |||
               |||
Qy      507 ProLeuAlaAlaAlaThrMetValArgIleGlyAlaGluHisValLysAlaSerAla 526
               |||
               |||
Db      238034 CCAGGTTTTCACAGACACGACCTTGAAACCTT-----GTTATGAAAGCAGCA 238081
               |||
               |||
Qy      527 MetGluAlaLeuGluLysValLysPheProSerArgTYrThrAspArgAsnSerAlaGluPro 546
               |||
               |||
Db      238082 CTTACAGCAGCTAGA-----GATGATAAATTTAATTGATGATG 238120
               |||
               |||
Qy      547 LysAspValGluGluThrValPheSerValAspProValGlyAsnAsnValGlyArgAsp 566
               |||
               |||
Db      238121 AAAGATTTTGAG-----TACGCTAAAGCAAAAGTT-----CTAATGCGTAAAGAG 238165
               |||
               |||
Qy      567 GlyLysProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsn 586
               |||
               |||
Db      238166 ---CGACGTAGTCTCATTTCTTCTGATGAAGAAATTAACACCTTCATGTAAGT 238222
               |||
               |||
Qy      587 AspTYrAlaIleArgAlaGlyValSerGluSerSerValAspGluThrThrValGluMet 606
               |||
               |||
Db      238223 GACACATGACTT----- 238234
               |||
               |||
Qy      607 SerValAlaAspMetLeuArgIleAla-----SerValLysIleLeuAlaAla 622
               |||
               |||
Db      238235 ---GTTGCACTTACTTCCAAAGACAGATCCAGTCCAGCAAAAGTAACTATTCCTCGA 238291
               |||
               |||
Qy      623 GlyValAlaIleGlyLeuIle-----SerLeuPheSerGln 634
               |||
               |||
Db      238292 GGTGCGTCTCTGCTGTAACAATGACATTAACCTGAAGCAGATAGACATAGTATTATCAAG 238351
               |||
               |||
Qy      635 LysTYrPheLeuLysSerSerSerPheGluArgLysArgAspMetValSerSerMetGlu 654
               |||
               |||
Db      238352 GAATATTTATTAATTAACCTGATGTTATTAGCTGTAGAGCAGCTGAAGAAATTAATA 238411
               |||
               |||
Qy      655 SerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeuProArgMet 674
               |||
               |||

```

Db 238412 TTGATACTATTACACAGGTGCA-----GGTAATGATATTGAACTGCAACCAATATG 238465
Qy 675 AspAlaArgThrAlaGluAenIlleValSerIleThr-----GlnIleValSer 691
Db 238466 -----GCAAGAAAAATGGTTGTGAAATGGGGCTTGAACCGCTTAATGACCA 238513
Qy 692 LeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMet 711
Db 238514 TTATTCATTGGT-----GAGCGTGAAGAGAGGTTTTATTTGGTAGAG 238558
Qy 712 LeuIleIleThrPheAspAlaAlaGluThrAlaGlnLeuGlyLeuValIleThrPheTyr 731
Db 238559 -----TGGGCACATCTTGAAGAACTTATGTAAGATACGGCTCGTATGTGTTGTAAGT 238609
Qy 732 ThrLeuLeuIleuSerValAspSer-----ValThrValSer 744
Db 238610 GAAGTTAAAAATTAATAGAAAGAGCTAGAGAAAATGTCAAGACATATTGACAAATAT 238669
Qy 745 AlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluIleuSerAlaCysLeuSerAsp 764
Db 238670 CTGTGATACGTTGCATGCTCTTGTCAACGACTACTTAAAGAGAAACATTACAGTAT 238729
Qy 765 LeuValHis-----ProGluAspAsnAla 772
Db 238730 GATATAGATCTTCTTATCAAGGGGAGAACTACCTCATTTATACCAAGAACTCTAAT 238789
Qy 773 ThrAspValArgThrIleThrThrArgTyrGluValPhe 785
Db 238790 ACTGAATTAATAAAAGAAATTCAAATATCTGAACCTTT 238828

RESULT 13
US-10-513-786-6
; Sequence 6, Application US/10513786
; Publication NO. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio
; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method for
; TITLE OF INVENTION: cereulide.
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513, 786
; CURRENT FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; PRIOR FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent version 3.1
; SEQ ID NO 6
; LENGTH: 11115
; TYPE: DNA
; ORGANISM: *Bacillus cereus*
US-10-513-786-6

Alignment Scores:
Pred. No.: 1.4 Length: 11115
Score: 111.50 Matches: 159
Percent Similarity: 33.7% Conservative: 117
Best Local Similarity: 19.4% Mismatches: 301
Query Match: 2.7% Indels: 241
DB: 8 Gaps: 36

US-10-600-070b-2 (1-801) x US-10-513-786-6 (1-11115)

Qy 105 IleArgArgAlaPheGluAlaArgValSerIleProGlnPheGlyPheSerAspAsp 124
Db 5389 ATTTCATGAATTTGTTGAAGCAGCTGTTCAACAAAACGCCA-----TCTGCACAA 5436
Qy 125 AlaLeuIleSerArg---ArgGlnIleLeuGlnAlaIaCysGluThrLeuSerAspPro 143
Db 5437 GCGCTTATTTATCGCAACAAAGATGATCATACGAGATGTTAAATGCCAGCGCAATATGCA 5496
Qy 144 ArgSerArgArgIleuThrAsnGluGlyLeuLeuAspAspGluAlaThrValIleThr 163
Db 144 ArgSerArgArgIleuThrAsnGluGlyLeuLeuAspAspGluAlaThrValIleThr 163

Db 5497 CTTCACATATAATTAAAGATGACGGTGTGACCAAAACCGAGTAGTGGCTGTATATG 5556
Qy 164 AspValProThrAspIleValProGlyAlaLeuCysValLeuGlnGluGlyGly----- 181
Db 5557 GATCGCTCTTTCGAGATGGTGTGGTATATTTAGCTGTTTAAAGACAGATGGTGGTAT 5616
Qy 182 -----GluThrGluIleValLeuArgValGlyGlu 191
Db 5617 TTGCAATTGATACAGCGTACCCGATGACAGAAACGAATACGCTCGCAAAATAGTAGAG 5676
Qy 192 Ala-----LeuLeuGlyGluArgLeuProIleSerPheIle 203
Db 5677 GCAACTATTCTCTTAACAAAGAGATGTAACCTTAAGAGACTTTA-----GATTTGAG 5730
Qy 204 GlnAspVal---ValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMet 222
Db 5731 GGGGAAGTTTTTACTTAATGATGATGACAGACTTTGAAGGGGATGGAAGATTTACAA 5790
Qy 223 AlaLeu-----Asp 225
Db 5791 AATATCAATATATCTACTAACCCTGCTTATATCATTTATACATCAGATCCACGGGAAT 5850
Qy 226 ProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeuIleuLeuGlnGlu 245
Db 5851 CCAAAAGTGTATGTATGATGACGCAATCAAGTGTGTGAATTTGCTACTCGAATTTACAAAG 5910
Qy 246 GluGlyAlaSerSerLeuAlaProAspLeuArgAla-----GlnIleAsp 260
Db 5911 AAA---TATCCGGTGTACGAGAAATTAAGCACTTGTAAAAACAATATACCTTTGAT 5967
Qy 261 GluThrLeuGluGluIle----- 266
Db 5968 GTTCTGTAGCCGAATTTTGGATGTTTCATGACAGTGCACACTTGTATATGCTGGA 6027
Qy 267 -----ThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAsp 281
Db 6028 CATGTGATGAAAAAGACCCAGAGAACTGATTAATGATTT----- 6069
Qy 282 AspTyrAlaAlaIleArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerVal 301
Db 6070 -----CAATGCCACAAAGGTACACATATTAACCTGTACATGATGCTA----- 6114
Qy 302 GlyGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluAsp-----Met 319
Db 6115 -----CATGCAATGTTACAGCGCTTGATGAAAGAAATTTTGCAATTATG 6159
Qy 320 AsnGluAlaPheLeuArgMetThrAlaAlaGluGlnVal----- 332
Db 6160 AATCGGTGAATAATATTTATCGTCGACAGAGAACTGTTTCAACAGAACTTTGTAAATGCA 6219
Qy 333 -----AspLeuPheValAlaThrProSerAsn 341
Db 6220 CTGTACGCTCATTTGTCAAATGTAATACTAGAAATCTTATATGGGCCAAACGAGAGAAAG 6279
Qy 342 IleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAlaPhe 361
Db 6280 ATTTAATGCCACAGCGTTTCTTATTCATTAAGAA-----ATGAATGTACTAATGACCG 6333
Qy 362 IleGlyIleLeuPheProIleLeuLeuGlnAspAlaAspGluPheGlnIleuGlnGln 381
Db 6334 ATTGAAAA-----CCACTTCTCAT 6354
Qy 382 AlaIleValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGlu 401
Db 6355 GTGGAACCTATATTTTGTAT-----CAAAACATCAAAAT 6390
Qy 402 IleAspPheGlyLeuGluArgIleuCys-----AlaLeuLeuIleGlyIle----- 417
Db 6391 GTACCAATGTGTGTACAGATGATGTGTCTGCGAGAAATATGTGTACGAAAAAGTTAT 6450
Qy 418 -----ValAspGluCysArgMetTrpLeuGlyLeuAspSerGluAsp 431
Db 6451 ATGAAGACCGGTGTAAACAGAAAGAAATTCGTGCAATCTTATGAACAAAGTGAA 6510

```

QY 432 SerGlnTyrArgAsnProAlaIleValGluPheVal-----LeuGluAsnSer 447
    |||||
DB 6511 AGAATGTACCGAACGGGTGATTGGTACGCTGTACCAATGGGAATTATTGAATTATTA 6570
QY 448 AsnArgAspAspAsnAsp-----AspLeuPro 456
    |||||
DB 6571 GGAAAGATAGATACCAAGTCAAGATAGAGAGCTTCGAATTGACCTTGGAATTGAA 6630
QY 457 GlyLeuCyLeuLeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAsp 476
    |||||
DB 6631 GCGGCATATTCCTCATTTAGAAAGTGTGTCAACCAATTGTT--ACACATATGACCGAT 6687
QY 477 ThrLeuAspLysLys-----PheLeuLeuLysAspTyrTyrAspAspPromet 492
    |||||
DB 6688 CATTAAGGTGCGAACCAAGATTGCGCATATGTTGTGAGCAAAAGTATGATGA----- 6741
QY 493 ValLeuSerTyrLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAla 512
    |||||
DB 6742 -----GACGAATTGCT----- 6753
QY 513 ThrMetAlaArgIleGlyAlaGluHisValLysSerAlaMetGlnAlaLeu----- 530
    |||||
DB 6754 -----GAACATGTGAAAAAGCGTTGCGCATATATATGTATACCA 6792
QY 531 -----GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
    |||||
DB 6793 AGTTATTTGCTTTCATGAAAGCATTCCTCTTAATAAAGAAAAAGTTGATGCGAAA 6852
QY 543 SerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGlyAsnAsn 562
    |||||
DB 6853 CAGTTCATTCGGTTGATCTTTATGAAACAGATATGATACATGCTGTGGGACCAAGA 6912
QY 563 ValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSerGluAsn 582
    |||||
DB 6913 AACGAAAAAGAACCAATGCTTTCGATT--ATTGGCAGAGCTTTGGGATATGAGAAAT 6969
QY 583 PheGluThrAsnAspTyrAlaIleArgAlaGly----- 593
    |||||
DB 6970 ATCAGTGTACATTAATTTCTTTAGCTTGTGTGTCATTCATTAATGCGACCAATTG 7029
QY 594 ValSerGlu-----SerSerValAspGluThr 602
    |||||
DB 7030 GTATCAAAAATTTATATGCTTTTCCAGTGAAGATGCTCTTAATAATGCTTTCACTAT 7089
QY 603 ThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu--Ala 621
    |||||
DB 7090 ACAACCTTAGCTACATGACGCGGTGTAGAAAGATTGTTGGTAAAGCGCTTGACGAA 7149
QY 622 AlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSerSer 641
    |||||
DB 7150 GTAGCTGTAAACGAGCGCATTCACAAAGATCTACCGAGAACATTAATTCGATTTGTCG 7209
QY 642 SerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIleGly 661
    |||||
DB 7210 TATTCACACAAAGATTTATTTCTTATCTACATGAGAAAGAAACCAATTACTATAT 7269
QY 662 SerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGluAsn 681
    |||||
DB 7270 ATCTTGTCGCTTGCGAT-----ATTATGGAACCTAGAT----- 7305
QY 682 IleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisAlaGlyLeu 701
    |||||
DB 7306 ---GTTACGCTATTTAAAAAGCAATCCAACTTAATGAAAGAAACCACTTCATTAAGT 7362
QY 702 MetLeuProGluValLeuAspGlyArgMetLeuLysIleTrpThrAspArgAlaAlaGlu 721
    |||||
DB 7363 GCAACATTGAAATCGTGAACGCGCAACCTGTGCAATC----- 7401
QY 722 ThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuLysSerValAspSerVal 741
    |||||
DB 7402 -----ATCCACGATGATATGAAATTCCTGTGCAATTTATTGACCTT 7443

```

```

QY 742 ThrValSerAlaAspGlyThrArg-----AlaLeuValGluAlaThrLeuGlnGluSer 759
    |||||
DB 7444 ACTGTATGCGACAGAGATTAACGATAGAAAGATGATATTAATTTAAAGACT 7503
QY 760 AlaCyLeuSerAspLeuValHis---ProGluAsnAsnAlaThrAspValArg 776
    |||||
DB 7504 AAAAGATATACATCTCCGAATATGTCCTTAATGATGTACAAATTGTTAAG 7557
RESULT 14
US-10-513-786-8
; Sequence 8, Application US/10513786
; Publication No. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio
; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method for
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513,786
; PRIOR FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 11115
; TYPE: RNA
; ORGANISM: Bacillus cereus
US-10-513-786-8
Alignment Scores:
Pred. No.: 1.4 Length: 11115
Score: 111.50 Matches: 159
Percent Similarity: 33.7% Conservative: 117
Best Local Similarity: 19.4% Mismatches: 301
Query Match: 2.7% Indels: 241
Gaps: 36
US-10-600-070b-2 (1-801) x US-10-513-786-8 (1-11115)
QY 105 IleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGlyPheSerAsp 124
    |||||
DB 5389 AUTCAGAAUUGUUGCAAGACAGCGUUCACAAAACCCA-----UUCGACAA 5436
QY 125 AlaLeuIleSerArg--ArgGlnIleLeuGlnAlaAlaCySerGluThrLeuSerAsnPro 143
    |||||
DB 5437 GCGCUUAUUUUAUGCAAAAGAGUAGCAUACGAGAGUUAUUGCGCAGCAAAUUGCA 5496
QY 144 ArgSerArgArgGluTyrArgGlnGlyLeuLeuAspAspGlnGluAlaThrValIleThr 163
    |||||
DB 5497 CUUGCAUAAUUAAGUUGCAAGUUGGACCAAAACCGAGUAGUUGCGGUUAUGU 5556
QY 164 AspValProTyrAspLysValProGlyAlaLeuCyValLeuGlnGlnGlyGly----- 181
    |||||
DB 5557 GAUCGCUUUCAGAGUUGUUGUUAUUAUGCGUUUUUUAAGACAGUGUGUGCUAU 5616
QY 182 -----GluThrGlnIleValLeuArgValGlyGlu 191
    |||||
DB 5617 UUGCCAUAUUGAUAACGCUAACCGAUGCAACGAAGAAUAGUCUUGCAAAUUAUGAG 5676
QY 192 Ala-----LeuLeuLysGlnArgLeuProLysSerPheLys 203
    |||||
DB 5677 GCAACUAUUCUUCUUAACAAAGAUUGUUAACCUUAAGAGUUCUUA-----GAUUUGAG 5730
QY 204 GlnAspVal---ValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMet 222
    |||||
DB 5731 GGGGAAGUUUUUAUCUUAAGUAGUAGCAACUGUUUGAAGGGGAUUGAAGAGUUUAACA 5790
QY 223 AlaLeu-----Asp 225
    |||||
DB 5791 AAUAUCAUAUAUUCUUAACUUGCUUAUUAUCAUUAUUAUUAUUAUUAUUAUUAUUAU 5850

```



```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3210
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3210

Alignment Scores:
Pred. No.: 0.347 Length: 3252
Score: 111.00 Matches: 164
Percent Similarity: 33.6% Conservative: 96
Best Local Similarity: 21.2% Mismatches: 250
Query Match: 2.7% Indels: 264
DB: 12 Gaps: 42

US-10-600-070b-2 (1-801) x US-11-136-527-3210 (1-3252)

QY 3 AlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCybArgLeuProProAla 22
DB 302 GCTGATGCTGTGAGCTGGCGAGTGAATCCATGAGGTGGCCATTGTCAGACCTGGA 361
QY 23 ThrThrIys-----LeuArgArgSerHisLeuSerHisThrThrIleCySerAla 39
DB 362 GTCGCCATGAGATTGTATTGAACAAAGAAATCTCGAGAAACTCCGGTACTGTGACC 421
QY 40 SerIleThrAlaAspArgLeuLeuSerAspPheAspPheThrSerAspSerSerSer 59
DB 422 TTGAAGAAACAAAGAGTTCTTAGTGAC-----AGTGCAGCT 460
QY 60 SerPheAlaThrAlaThrAlaThrAlaThrLeu-----ValSerLeuProPro 75
DB 461 GGCATGGCCATCAAGAACCAAGAGCTACGCTGATTTCAGACACTCTTGGAAG 520
QY 76 SerIleAspArgProGlnArgHisValProIleProIleAspPheThrGlnValLeuGly 95
DB 521 CAGCAGATTAACTT-----CATGTGCTCTT-----TACCAGTCCCTGTTCC 562
QY 96 AlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIys 115
DB 563 CCGAACAATGAGCTCAATGTTGACCCACAGAGCAGACTGTGGCTTCCAGATCAGT-- 619
QY 116 ProProGlnPheGlyPheSer-----AspPheAlaLeuIleSer----- 122
DB 620 ---CCGACGCTGAGTTCTCTCCGAGAGAGTGTGGGATTCCTCAACTACCTCCCT 676
QY 123 -----AspPheAlaLeuIleSer----- 128
DB 677 TCCCTGGCTGAAGATTTTGCAAGAACACTATTAAAGATGACAGTATCACCTGCGCAGCC 736
QY 129 -----ArgArgGlnIleLeuGlnAlaAla----- 136
DB 737 TTTTTCACCAAGCCCGACGCCGAGCTGTCTGACAGCTGTCTGTATGCTGCTGACCTC 796
QY 137 -----CysGlnThrLeuSerAspProArgSerArgArg 147
DB 797 GTGCTGAGCTCATCAATGACAACTGCGCAGCCCTCAGCTATGCTGCTTCCGCCG 856
QY 148 Glu-----TyrAsnGlnGlyLeuLeuAspAsp 156
DB 857 AAAGATATCAATTCATGTCAGACAAATATCATGTTTCATGACATGAGC----- 904
QY 157 GlnGlnAlaThrValIleThrAspValProTyrAspValProGlyAlaLeuCybAla 176
DB 905 TCGGAGCAGCACTGTGTGTCATCTGACCTTACCAACGCTGAG-----ACT 952
QY 177 LeuGlnGlnGlyIleGlnThrGlnIleValLeuArg---ValGly----- 190
DB 953 AAGGAGCTGTGGAGCAGCAGCAAGTACAGATCCGGGGGTGGGATTGACGCGACCTGG 1012
QY 191 -----GlnAlaLeuLeuGlyGlnArgLeuProIysSerPheIysGlnAsp 205
DB 1013 GGTGCTGTGAAGATGAGCTTCTGGCTGCGAGACCACTGGCTAAGCTTTCATGAGCAG 1072

```

```

QY 206 ValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAsp 225
DB 1073 CGCAAGGCGCAAGAAAGCC-----AAGGATGTTCCGGAAGAAACCCCGAGCCATG--- 1120
QY 226 ProProAspPheIleThrGlyTyrGlnPheValGlnGlnAlaLeuIysLeuGlnGln 245
DB 1121 -----GCCAAACTGCTTCGGGAA 1138
QY 246 -----GlnGlyAlaSerSerLeuAlaProAspLeuArgIleAspGln 261
DB 1139 GCCAATCGGCTTAAACCGTCTGAGTCCCAATGCTGATCACTGGCAGATGTAAGGC 1198
QY 262 ThrLeuGlnGlnIle-----ThrProArgTyrValLeuGlnLeuGly 276
DB 1199 TTGATGAGAGATGTGAGCTTCAAGSCAAAGTAAGTCACTGAGTGAAGTGTGTGT 1258
QY 277 LeuProLeuGlnIleAspAspTyrAlaAlaIleArgLeuAsnGly----- 290
DB 1259 GCAGATTGTTGTAT-----CGAGTCCCTGGCTGTGACAGAGCCCTG 1303
QY 291 -----LeuSerGlyValArgAsnIleLeuThrPheValGlyGly 304
DB 1304 CAGAGTCTGAGATGAGCTGTGATCAAAATTGACAGAGTGATC----- 1345
QY 305 GlyAlaSerAlaLeuValGlyGlyLeuThrArgGlnIlePheMetAsnGlnAlaPheLeu 324
DB 1346 -----CTGTGGTGGGCGCCACTCGTCTTCCCAATTCAGAGGCTGTGCTG 1393
QY 325 ArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSerAsnIleProAla 344
DB 1394 AACCTGTGGGCAAGAGAGAACTAGCA-----AAGAACTCAATGTCC 1435
QY 345 GluSerPhe-----GluValTyrGlnValAlaLeuAlaGlnAla 360
DB 1436 GATGAAGCAGCTGCCATGGGGCGCTGTACAGCA-----GGGCACTGACCAAGCC 1489
QY 361 PheIleGlyIleIysProHisLeuLeuGlnAspAlaAspIysGlnPheGlnGln 380
DB 1490 TTC---AAAGTGAAGCACTTGTGTGCGCGAT----- 1519
QY 381 GlnAlaIleValMetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAspTyr 400
DB 1520 -----GCTGTTATTACCCCATCTGTGTGAGTTCAAGAGGAGGTGAG 1564
QY 401 GlnIleAspPheGlyLeuGlnArgGlyLeuCybAlaLeuLeuIleGlyIysValAspGln 420
DB 1565 GAG-----GAGCTGTGGCTTCAAGAGCTGAAAGCAATATA----- 1600
QY 421 CysArgMetTyrLeuGlyLeuAspSerGlnAspSerGlnTyrArgAsnProAlaIleVal 440
DB 1601 ---GCTGTG-----CTCTTCTCCGAAATGGGGCCTTACCCTCAGCGCAAAATGTCATC 1648
QY 441 GluPheValIleGlnIleAsnSerAsnArgAspAspAsp----- 453
DB 1649 ACCTTT-----AACGATACAGCATGATTTCAACTTTCACATCACTAC 1693
QY 454 -----AspLeuPro-----GlyLeuCybIleLeu 462
DB 1694 GGTGACCTGGGCTTCTGGGGCGCTGAGATCTTCGGATTTTGGCTCCGAGATCTGACC 1753
QY 463 GluThrTyrPheAlaGlyVal-----ValPheProArgPheArgAspThrIysAspIys 480
DB 1754 ACAGTGAACCTTAAAGGTGTGGAGAGAGACTTCAAGAAATATCTCTATGATGCTCAAA 1813
QY 481 IysPheIysLeuGlyAspTyrTyrAspAspProMetValIleuSerTyrLeuGlnArgVal 500
DB 1814 GGCATCAGGCCCATTTAACTTAAGTACGAGAGTGTCTCAGT---TTAGCAGGGTG 1870
QY 501 Glu-----ValValGlnGlySerProLeuAlaAlaAlaThrMetAla 515
DB 1871 GAGTCTGATTGCAAGACCTGTGTGAGGACAGCCCA---GAGAAAGAGCTTCTTACC 1927
QY 516 ArgIleGly-----AlaGlnIleValIysAla 524

```



```
Db      1928 AAACCTGGCAACACATTCCAGCCCTGTTGGCGGCTGATACCTCATCAGATGCCAAAGAG 1987
      :::::||||
Qy      525 SerAlaMetGlnAlaLeuGln-----LysValPheProSerArgTyrThrAspArgAsn 542
      :::::|||||
Db      1988 AATGCTACTGATGCTGTACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAG 2047
      AATGCTACTGATGCTGTACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAG
Qy      543 SerAlaGluProLysAspValGlnGluThrValPheSer----- 555
      |||||
Db      2048 CCTGCAGAACAGGGGAGACTCAAGAGAGAACTGAAGCCCAATGAGAGATACCTCCAG 2107
      |||||
Qy      556 -----ValAspProValGlyAsnAsnValGlyArgAspGlyGlu----- 568
      ::|||
Db      2108 CCTCCACCTCTGAGGCTTAAGGG---GATGCAGCCCGTGTAGGGAGAGAAACACCTGATGAA 2164
      ::|||
Qy      569 ---ProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
      ::|||
Db      2165 AAAGAAAGTGGGAGCAAGCTGTAGGCCCAATGAGAAAGGGGCGAGCGAGGGGCT 2224
      ::|||
Qy      588 TyrAlaIleArgAlaGlyValSerGluSerSer----- 598
      ::|||
Db      2225 GAGGCTGTCCCTCCAGCTCCCGAGAGAAAGAAAGACAGAACTGCCCGGAAACAGAA 2284
      ::|||
Qy      599 ---ValAspGluThrThrValGluMetSerValAlaAspMet 611
      |||||
Db      2285 ATGCTGAGAGAGATAGTGTGAACTGCTGTCTTGGAACCTG 2326
      |||||
```

Search completed: February 21, 2006, 14:51:15
Job time : 3423 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 13:54:58 : Search time 6309 Seconds
(without alignments)
17842.720 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagctcgtgacgct.....gctctgtctgtgacataa 2406

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	902.2	37.5	919	4	CNS09YJH
2	633	26.3	741	4	CNS09YJH
3	633	26.3	741	4	CNS09YJH
4	482	20.0	561	1	AI998415
5	458	19.0	534	5	B0834167
6	447	18.6	2307	10	CL965374
7	436	18.1	833	7	CO071968
8	429.4	17.8	897	6	CD572714
9	360.8	15.0	785	6	B2437564
10	359.2	14.9	624	8	CX195205
11	344.2	14.3	769	7	CO075595
12	335.4	13.9	741	8	CX543653
13	328.8	13.7	631	5	BU046755
14	326.6	13.6	703	3	CN918842
15	321	13.1	415	3	BP785511
16	316	13.1	611	8	BP782886
17	306.8	12.8	611	8	CX171798
18	304.4	12.7	341	9	AY199896
19	300.6	12.5	728	8	CX674370
20	299.2	12.4	434	3	BP620404
21	293.4	12.2	832	7	CO079829
22	288.2	12.0	697	8	CX173275

C 23	287.6	12.0	821	7	CK090561
C 24	287.6	12.0	922	7	CV264112
C 25	286.6	11.9	819	7	CV241483
C 26	283	11.8	898	8	DR830109
C 27	280.2	11.6	722	7	COL17046
C 28	278.8	11.6	660	2	B1268376
C 29	272.2	11.3	839	7	CK936162
C 30	270.4	11.2	866	6	CD573715
C 31	270.2	11.2	723	7	CO079828
C 32	265	11.0	746	8	CX674369
C 33	261.8	10.9	768	8	CX543654
C 34	260.2	10.8	574	1	AM696905
C 35	257.4	10.7	593	7	CN904734
C 36	254	10.6	849	7	COL17047
C 37	249.6	10.4	871	10	DU050213
C 38	235	9.8	820	8	DR972351
C 39	231.8	9.6	751	11	CR486014
C 40	226.6	9.4	521	7	CN914655
C 41	226	9.3	683	7	CO076192
C 42	224	9.3	537	2	BE490117
C 43	220	9.1	501	5	BU049629
C 44	218.6	9.1	508	3	BQ048486
C 45	214.8	8.9	451	1	AJ802091

ALIGNMENTS

RESULT 1
CNS09YJH
LOCUS
DEFINITION
CNS09YJH 919 bp mRNA linear HTC 04-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTS892C08 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION
BX841670.1 GI:42406830
VERSION
HNC: GSUT CDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM

REFERENCE
Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE

Unpublished
2 (bases 1 to 919)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
UNGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information Center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_FR/Full length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. 919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

Db 121 TTCAGCCAGAGTATTTCTTAAAGAGCTCATCTTTCAAGCAGCAAGATATGTTCT 180
 Oy 1954 TCTATGGAATCTGATGCGTACCATGCGGTGCTGAGCTGACGATGCAATTAAGCACT 2013
 Db 181 TCTATGGAATCTGATGCGTACCATGCGGTGCTGAGCTGACGATGCAATTAAGCACT 240
 Oy 2014 CCCAGATGAGTCTAGAGCTGAGAGATATAGTATCCAAAGTGGCAGAAATTAAGTCT 2073
 Db 241 CCCAGATGAGTCTAGAGCTGAGAGATATAGTATCCAAAGTGGCAGAAATTAAGTCT 300
 Oy 2074 CTGGCTTTTGGGCTGATGACCCGATAGAAATGTTACCAAGGTTTGGATGGCGAATG 2133
 Db 301 CTGGCTTTTGGGCTGATGACCCGATAGAAATGTTACCAAGGTTTGGATGGCGAATG 360
 Oy 2134 CTGAATTTTGGATGACAGAGAGCTGAAATGCGGAGTTTGGTGTATGATAT 2193
 Db 361 CTGAATTTTGGATGACAGAGAGCTGAAATGCGGAGTTTGGTGTATGATAT 420
 Oy 2194 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCGTGTCTG 2253
 Db 421 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCGTGTCTG 480
 Oy 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGTTATCTGATTTGGTTCATCCAGAAACAAT 2313
 Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGTTATCTGATTTGGTTCATCCAGAAACAAT 540
 Oy 2314 GCTACTGATGTCGAACCTTACACAAAGATAGAAATTTTGTGTTCAAGTCAAGGTGG 2373
 Db 541 GCTACTGATGTCGAACCTTACACAAAGATAGAAATTTTGTGTTCAAGTCAAGGTGG 600
 Oy 2374 AAATACATGAGAGCTCTGTTTGCATCATTA 2406
 Db 601 AAATACATGAGAGCTCTGTTTGCATCATTA 633

RESULT 3
 CENS09Z55 741 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSI751J12E09 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).

ACCESSION BX833051.1 GI:42455503
 VERSION
 KEYWORDS HTC; GSUT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 741)
 Castelli V., Aurty J.M., Jailion O., Wincker P., Clepet C.,
 Menard M., Cruaud C., Quetier F., Scarpetti C., Schachter V.,
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation

TITLE Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSI751J12E09 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress)

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 741)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 101 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aurty J.M., Jailion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_Full1
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 location/Qualifiers

FEATURES
 source
 1..741
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSI751J12E09"
 /release_type="Silique"
 /ecolyse="Col-0"
 /plasmid="pCMVSPORT_6"
 1..741
 /gene="At5g42480"

ORIGIN

Query Match 26.3%; Score 633; DB 4; Length 741;
 Best Local Similarity 100.0%; Pred. No. 1.1e-167;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1774 GCTGGGGTCTCAGAGTAGCGTTGATGAAACTGATGTTGAAATGTCCTGCTGATATG 1833
 Db 1 GCTGGGGTCTCAGAGTAGCGTTGATGAAACTGATGTTGAAATGTCCTGCTGATATG 60
 Oy 1834 TTAAAGAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTGGAATTTCACTG 1893
 Db 61 TTAAAGAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTGGAATTTCACTG 120
 Oy 1894 TTGAGCGAAGATTTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTCT 1953
 Db 121 TTCAGCAGAGATTTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTCT 180
 Oy 1954 TCTATGGAATCTGATGCGTACCATGCGGTGCTGAGCTGACGATGCAATTAAGCACT 2013
 Db 181 TCTATGGAATCTGATGCGTACCATGCGGTGCTGAGCTGACGATGCAATTAAGCACT 240
 Oy 2014 CCCAGATGAGTCTAGAGCTGAGAGATATAGTATCCAAAGTGGCAGAAATTAAGTCT 2073
 Db 241 CCCAGATGAGTCTAGAGCTGAGAGATATAGTATCCAAAGTGGCAGAAATTAAGTCT 300
 Oy 2074 CTGGCTTTTGGGCTGATGACCCGATAGAAATGTTACCAAGGTTTGGATGGCGAATG 2133
 Db 301 CTGGCTTTTGGGCTGATGACCCGATAGAAATGTTACCAAGGTTTGGATGGCGAATG 360
 Oy 2134 CTGAATTTTGGATGACAGAGAGCTGAAATGCGGAGTTTGGTGTATGATAT 2193
 Db 361 CTGAATTTTGGATGACAGAGAGCTGAAATGCGGAGTTTGGTGTATGATAT 420
 Oy 2194 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCGTGTCTG 2253
 Db 421 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCGTGTCTG 480
 Oy 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGTTATCTGATTTGGTTCATCCAGAAACAAT 2313
 Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGTTATCTGATTTGGTTCATCCAGAAACAAT 540
 Oy 2314 GCTACTGATGTCGAACCTTACACAAAGATAGAAATTTTGTGTTCAAGTCAAGGTGG 2373
 Db 541 GCTACTGATGTCGAACCTTACACAAAGATAGAAATTTTGTGTTCAAGTCAAGGTGG 600
 Oy 2374 AAATACATGAGAGCTCTGTTTGCATCATTA 2406
 Db 601 AAATACATGAGAGCTCTGTTTGCATCATTA 633

RESULT 4
 A1998415 561 bp mRNA linear EST 08-SEP-1999
 LOCUS 701545606 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
 DEFINITION thaliana cDNA clone 701545606, mRNA sequence.
 ACCESSION A1998415
 VERSION A1998415.1 GI:5845320
 KEYWORDS EST.

QY 2141 TTGGACCTGACAGACAGCTGAACCTGCGAGCTTGGGTTGGTTATGATTTATACACTGT 2200
| | | | |
Db 242 TTGGACCTGACAGACAGCTGTTGGAACCTGGAGCTTGGGCTGGTTATGATTTATACACTGT 301
| | | | |
QY 2201 TGAACCTATCTGTTGACAGTGTGACAGTCTCAGACAGATGGAACCCGCTCTGGTGAAG 2260
| | | | |
Db 302 TGAACCTATCTGTTGACAGTGTGACAGTCTCAGACAGATGGAACCCGCTCTGGTGAAG 361
| | | | |
QY 2261 CAACCTGAGAGAGTCTGTTGATCTATCTGATTTGGTTCATCCGAAACAAATGCTACTG 2320
| | | | |
Db 362 CAACCTGAGAGAGTCTGTTGATCTATCTGATTTGGTTCATCCGAAACAAATGCTACTGAG 421
| | | | |
QY 2321 ATGTCAAGACCTACACACAGATACGAAGTTTCTGCTCAGTCAAGTGAAGAAATCA 2380
| | | | |
Db 422 ATGTCAAGACCTACACACAGATACGAAGTCTGCTCAGTCAAGTGAAGAAATCA 481
| | | | |
QY 2381 CTGAAGGCTCTGTTCTTGCATCATTA 2406
| | | | |
Db 482 CTGAAGGCTCTGTTCTTGCATCATTA 507
| | | | |

RESULT 6
CL965374 2307 bp DNA linear GSS 21-SEP-2004
LOCUS OeIFCC012338 Oryza sativa Expressed Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
CL965374
VERSION CL965374.1 GI:528385433
SOURCE GSS.
ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriothalideae; Oryzaceae; Oryza.
1 (bases 1 to 2307)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, D., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL Contact: Chen Chen
COMMENT Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
Source location/Qualifiers
1..2307
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_11b="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

Query Match 18.6%; Score 447; DB 10; Length 2307;
Best Local Similarity 53.7%; Pred. No. 1e-114;
Matches 1256; Conservative 0; Mismatches 880; Indels 201; Gaps 8;

QY 108 CTGCTCCGACGAGAAATGGCGGACCGCTTCTCTCGACTTAATTCACCTCCGATTC 167
| | | | |
Db 132 CTGGGCGGAAAGCGCTCTTCCGCGACTTCACCTCTCCACCGCGCGCCCTCCGACCC 191
| | | | |
QY 168 CTCTCTCTCTCTCTCTGCGGACCGGACACACACGCACTCTGCTCTCTGACACATC 227
| | | | |
Db 192 GCGGTCCCGGCTT 251
| | | | |
QY 228 TATGATCTGTCGGAAGCGGACGTCCTCCATCCCATGATTTCTACAGGTAAT----- 280
| | | | |

Db 252 CCCGAGCCGCGGACGCTCCCTCCGCTCAAGTGAATTTCTACAAAGGTGAGGCCCG 311
| | | | |
QY 281 -----TAGAGCTCAACACATTTCTTAACGATGGAATTCAGAAAGC 323
| | | | |
Db 312 TGAATCTGCTGCGGTTCTAGAGGCGAGGCCACATTTCTTGGCGATGCAATCAGAGGCG 371
| | | | |
QY 324 ATTCAAGCTAGGATTTGAAACCGCGCAATTCGATTTCAAGCAGCAGCTTTATCAAG 383
| | | | |
Db 372 GTTGAGGCGACGGAATACCAAGCAACCGGATATGCTTACAGCAGCATGCTCTGTTGG 431
| | | | |
QY 384 CCGAGACAGATTTCTTAACCTGTTGCGAACTCTGCTTAATCTCGCTCTAGAAAGA 443
| | | | |
Db 432 TCGTCGCAAAATCTGAGATGCGCCATGACCTCTCATGAACAGAAATCCCGCATCTCA 491
| | | | |
QY 444 GTTCAATGAAGGTCTTCTGATGATGAAGAGCTACGATCATGATGCTCTTGGGA 503
| | | | |
Db 492 GTATGATCTGCGCTTCTGAGAACCGTGAAGAGCTCTACCATGATATGCTTGGGA 551
| | | | |
QY 504 TAAAGTTCTGCGGAGCTCTGCTGATTTGAAGAGGTGTGAGACTGAGATGATTTCTCG 563
| | | | |
Db 552 CAAGAGGCTCGG-----GAGGCACTTGCCTGCTTGT 584
| | | | |
QY 564 GATTGTGAGGCTCTGCTTAAGAGAGGTTCCTAAGTCTTTAAGCAAGATGTGTTT 623
| | | | |
Db 585 AACTGAGAAACAGTTGCTCTGATCGGCCAACCCAGCGCTTCAAGCAGGACGCTGTCT 644
| | | | |
QY 624 AGTTATGCGGCTTCCGTTCTCATGCTCTGAGAGGATGCTATGAGCAATGATCCACTGA 683
| | | | |
Db 645 AGGATGAGCTCTGCTTATGATGATCATCAAGGAGATGCTATGAGCAAGCCCTCCAGA 704
| | | | |
QY 684 TTTTATTAATCTGTTATGAGTTGTTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGC 743
| | | | |
Db 705 TGTATTTGCTGCTGCGAGGCTGCTCAGAGAGGCTCTCAAGCTCTTGAAGAAATGAGAC 764
| | | | |
QY 744 AAGTACCTTGAACCGGATTTACGTCACAAATTTGATGATGACTTTGAGAGATCACTCC 803
| | | | |
Db 765 AAGCAATCTGCAACCTGATCTGCTTCAAGATTTGAATGAATCTCGAGAGAGATTAACCC 824
| | | | |
QY 804 GCGTATGCTTGTGAGTACTTGGCTTACCGCTTGGTGAATTTACGCTCCGAAAGACT 863
| | | | |
Db 825 TCCTGTGATTTGAGAGCTTCTCCCTCTATTTGACACAGACATCATTAAGAGCCCA 884
| | | | |
QY 864 AATGTTTAAAGCGGTGCGGAAATTTTGTGATCTGTTGAGAGAGGTGAGATCAGC 923
| | | | |
Db 885 AGAAGGCTTCAAGGTGCGAGAAACATTTGTGAGAGCTTGGCAGAGAGATTTGCTTAC 944
| | | | |
QY 924 TCTTGTGGGGATTTGACCCGAGAGATTTATGAATGAGCGTTTATGAAATGACAGC 983
| | | | |
Db 945 CGTTGAGAGAGATTTCTGTAAGCCTTCAAGAGAGGCTTTTGTGAGATGACATC 1004
| | | | |
QY 984 TGCTGAGCAGTGTGATCTTTTGTAGTACCCCAAGCAATTTCCAGCAGATCTTGA 1043
| | | | |
Db 1005 AATTTGAACAGATGATTTCTTTTCAAAAACAGCAATGATTCCTCGAAGGTTTGA 1064
| | | | |
QY 1044 AGTTTCAAGATTTGACCTTGTGCTGCTGAGGCTTTTATGTTGTAAGAGCAACCT 1103
| | | | |
Db 1065 AATTTCAATGTTGACCTTGTGACATGCTCTCAAGCAATTTATGATTAAGGCAACCAAT 1124
| | | | |
QY 1104 TTTACAGATGCTGATTAAGCAATTTCAAGCACTTCAAGCGCTAAGGTAATGCTATGA 1163
| | | | |
Db 1125 CATCATGATGAGGAGATGATCTTTTGAACACCTCAAGAGTTCA----- 1169
| | | | |
QY 1164 GATTCTGCGATTTGTTATGATACAGCAATTAATGGAGATGACTTGGCTTGAAG 1223
| | | | |
Db 1170 -----CATAGGTTCTCATTTATGTTATGATGATGATGACCTTGCATTTGAAG 1220
| | | | |
QY 1224 GGGACTGTGACCTGTTATAGCAAAAGTGTGATGCGTATGAGTTGGGCTTTAGA 1283
| | | | |
Db 1221 GGCATTTCTGCTATTGTTATGTCGAAATTTAGCAAGTGAATGCTTGGAAATGA 1280
| | | | |
QY 1284 CAGTGAAGATTTCAATATAGAAATCAAGCTATTTGAGATTTGTTTGAAGATTC--- 1340
| | | | |
Db 1281 TATGATGTTTCAACATAGAGAACCCCAAAATTTCTAGATTTATGTTGACCAATCTAG 1340
| | | | |

1341 AAATCGTATGACATGATGATCTCCTGGACATATGCAAAATGTTGGAAAACCTGGTGGC 1400
 1341 CATCAGTGAAGAAATATCTCTTCCAGGGCTGTGCAACTTTTGGAGCTTGCTAT 1400
 1401 AGGGGTGTCTTCTCTAGTTCAGAGACCAAAAGATTAATACTTCCGGGAGTA 1460
 1401 CTTTGAAGTCTTCTCTAGAGACAGATACCTGGGGCATAGTTCAGACTTGGAGATTA 1460
 1461 CTATGATGATCTATGTTTGGTCTCTTGGAAAGTGGAGTACTTCAAGGTTCTCC 1520
 1461 CTAGAGATGATCCAGAAATTTTAACTTCTAGAAAGATGAGGGTGTGTCTTCA 1520
 1521 TTTAGCTGCTGCTCACTATGAGCAAGATTTGAGCCGAGCATGTGAAGCTGCTAT 1580
 1521 TTTGGCTGCTGCTGCTCTATGCAAACTTGGTCTCAAGCTACAGCTGACTGGTA 1579
 1581 GCAAGGCACTGACAGAAATTTTCTTCCCTACATACAGATGAAAATCGGCTGAACCA 1640
 1580 -----CTGTGAATCAAA 1592
 1641 GATATGACAGACAGACTGTTTATGATGATCTGTTGTTAACAATGAGCCGTATG 1700
 1593 TGCTATTCAAGCG----- 1605
 1701 TGAGCTGTGTCTTTTATGCAAGAGCTGTAGACCTCTGAATACTTTGAACCTATATGA 1760
 1606 -----TTCAACAAGGTTTTTCCATGATAGAACAGTTAGACAGCTAGC 1649
 1761 TTATGCAATTCGAGCTGGGCTCTCAGAGAGTACGCTGTATGAAAATCTATGAAATGTC 1820
 1650 CATGAAAAATCTAAAGATGGCCCTGGGGGATATCTTGAAAATTTTGACCAAGAAAATGC 1709
 1821 CGTTGCTGATATGTTAAGAGAGCAATGTAAGATCTTACGCTGCTGTGTGGAAATGG 1880
 1710 ACCGCTCATGATTCGAAAATGCGGCTGGAATATCTGCTGCGCCACTGTTTGC 1769
 1881 ACTGATTTCACTGTTCAAGCAGAAATTTTCTTAAAGACGCTCATCTTTTCAACGCA 1940
 1770 ACTGTTGGCAGTATTTGGGCGCAAAATTTTCCCTGCTGTAAGAGGCCCTTCTGTATAG 1829
 1941 GATATGATTTCTTCTATGAAATCTGAT-----GTGCTACCATAGGGTCACTGAC 1991
 1830 GAGTGAACATGATCTGTGGCAGTGTGCTAATAGTGTGATCTGATGATCTGTGACT 1889
 1992 AGCTGACGATTCGAAACACCTTCCCAATGAGATGCTAGACCTGACAGAAATATGATATC 2051
 1890 AGATGAAGATCCAGTACATATCTTGAATGATGCGAAGCTGGCAGAAATATGTTGC 1949
 2052 CAAGTGGCAGAAATTAAGTCTCTGAGCTTTTGGGCTGATACACGCAATGAAATGTTAC 2111
 1950 CAAGTGGCAGAAATTAAGTCTCTGAGCTTTTGGGCTGATACACGCAATGAAATGTTAC 2009
 2112 AGAGTCTTGGATGGGGAATGCTGAATTTGAGCTGACAGACAGCTGAAATGTCGCA 2171
 2010 AGAGTCTTGGATGGGGAATGCTGAATTTGAGCTGACAGACAGCTGAAATGTCGCA 2069
 2172 GCTTGGTGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2231
 2070 TCAATGGTGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2129
 2232 AGCAGATGGAACCGTCTGCTGAGGAGCACTCTGAGAGAGTCTGTTCTATCTGA 2291
 2130 CTTAGATGCTGACAGAGCACTGAGGCTGAGATGATGAGGCAAGCCCACTTACTGA 2189
 2292 TTTGGTTCATCCAGAAAAACAATGCTATGATGTCAGAACCTTACAAACAATGCAAGT 2351
 2190 TGTATGAGAGCCCAAGAAACAATGATTAATGACAAAAATACACTACCGGATATGAGAT 2249
 2352 TTTCTGTCCAG---TCAGGGTGAATACTGAAGGCTCTTCTTGTGATATA 2405
 2250 GGCCTTCTCCAGTGAAGGGTGAAGATATCGAAGAGAGCACTCTCAAGTCTGA 2306

RESULT 7
 CO071968 813 bp mRNA linear EST 15-JUN-2004
 LOCUS GR_Ea30N03.r GR_Ea Gossypium raimondii cDNA clone GR_Ea30N03 3',
 DEFINITION mRNA sequence.
 ACCESSION CO071968
 VERSION CO071968.1 GI:48741449
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 813)
 REFERENCE Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
 Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and
 Wing, R. A.
 TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 plate: 30 row: N column: 03.
 FEATURES
 source location/Qualifiers
 1..813
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Ea30N03"
 /ribose_type="whole seedlings"
 /dev_stage="first true leaves"
 /lab_host="PH10B"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV. Library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Colonies
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
 Query Match 18.1%; Score 436; DB 7; Length 813;
 Best Local Similarity 72.3%; Pred. No. 9,7e-112;
 Matches 584; Conservative 0; Mismatches 215; Indels 9; Gaps 1;

107 TCTGCTCCGCGCAGAAATGAGCGGCTCTTCTTCCGACTTCAATTTCACCTCG-- 163
 5 TCTGCTCCGCGCAGTAATGAGCGGCTCTTCTTCCGACTTCAATTTCACCTCGCC 64
 164 -----ATTCTCTCTCTCTCTCTCTTGGCCACCGCCACCAACCGGCACTCTGCTCTC 217
 65 CCGATTAATTCGCTCTCTCTCTCTCTTCTTCCGACCGCACTCTTCTCCCTTACCTC 124
 218 TGGCAGCATATTAATGATCGCCGAGCGCAAGTCCCATCCCATTTGATTTTTCACAG 277
 125 CCTTCTCGCCCTTCTCTCTCCGAGCGCAAGCTTTCATTTCCCTTGAATTTTCAAG 184
 278 TATTAAGAGCTCAACACATTTCTTAACCGATGAATCAAGAGCATTAAGAGTGAAG 337
 185 TTTTGAAGCGGAGTCAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATG 244
 338 TTTGGAACCGCGCAATTCGTTTCAAGGAGCAAGCTTTAATCAACCGGAGCAATTC 397
 245 TTTGGAACCGCGCTCAATTCGTTTCAAGGAGCAAGCTTTAATCAACCGGAGCAATTC 304
 398 TTTCAAGCTGCTTGAAGCAATTCGTTTCAAGGAGCAAGCTTTAATCAACCGGAGCAATTC 457
 305 TTTTACTGCTCTGTAAGCAATTCGTTTCAAGGAGCAAGCTTTAATCAACCGGAGCAATTC 364
 458 TTTCTGATGATGAAGAGCTACAGTCACTGATGATGATGATGATGATGATGATGATGATG 517

```

Db      365  TTGCGACGATGACGTGACCTATCATCACTCAAGTCCCTGGGACAAAGGTTCTCTGAG 424
Qy      518  CTCTCTGTATTTGCAAGAAAGGTGTGAGACTGAGATAGTTCTTGGGTTGTAGAGCTC 577
Db      425  CATTTGCGGTGTTGCAAGAAAGCTGGGAGACGTGAAGTGTGTCTTCAATTGGGAGAGTT 484
Qy      578  TGTCTTAAGAGAGGTTGCTTAAGTGTGTCTTAAAGAGAGTGTGTCTTAAAGTGTGTG 637
Db      485  TGTCTGAAGAGAGGTTGCTTCAAGGCAATTTAAAGAGAGTGTGTCTTCAAGTGTGTG 544
Qy      638  CGTTTCTGATGCTGTGAGGAGTCTATGCGATTGATCCACTGATTATTAATGCTGTT 697
Db      545  CTTATGTGACTGTTGCAAGGAGTCTATGCTTGTGATCTCCAGATTTCATGTTGTTG 604
Qy      698  ATGAGTTTGTGAGGAAGCTTTGAAGCTTTTACGAGAGAGAGAGCAAGTACCTTGAC 757
Db      605  GTGAGGTGCTTGAAGAGGCTTTGAAGCTGTGTGACGAGAGAGAGTCCAGACGTCTCTC 664
Qy      758  CGGATTTACGTGCAAAATTGATGAGCTTTGGAAGAGATCACTCCGCTTATGCTTGG 817
Db      665  CGGATTTGCAAGTCAAGATTGATGAGCAATTAGAGAGATTACCCACGTTGTCTTGG 724
Qy      818  AGCTACTTGGCTTACCGCTTGTGATGATTAACGCTGCAAAAGACTAAATGTATTAGCG 877
Db      725  AACTTGTAGCTTGGCCCTTGTGATGATGATGCGAAGAGAGAGGAGGCTTTATG 784
Qy      878  GTGTGCGGAATATTTGTGTCTGTTGG 905
Db      785  GTGTAGCGCAACATATATATGAGCTGTTGG 812

```

```

RESULT 8
CD573714      897 bp      mRNA      linear      EST 12-JUN-2003
DEFINITION   UCRPT01_01_P12_T3 Poncirus trifoliata CTV-challenged cDNA library -
ACCESSION    CD573714
VERSION      CD573714.1 GI:31669616
KEYWORDS     EST.
SOURCE       Poncirus trifoliata
ORGANISM     Poncirus trifoliata
Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 897)
Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D.,
Mananaker,S., Choi,Y. and Kingan,T.
Development of EST Resources and New Genetic Markers for California
Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
Unpublished (2003)
Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 9097874137
Fax: 9097874437
Email: mikeal.roose@ucr.edu
Seq primer: T3.
Location/Qualifiers
1. 897
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_01_P12"
/tissue_type="Phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TUC121"
/clone_1lb="Poncirus trifoliata CTV-challenged cDNA
library - UCR"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse at University of California, Riverside. The

```

ORIGIN

```

Query Match 17.8%; Score 429.4; DB 6; Length 897;
Best Local Similarity 70.4%; Pred. No. 7.5e-110;
Matches 574; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

scion was a open-pollinated (very probably selfed) seedling of Poncirus trifoliata cv Pomeroy that was selected as homozygous for the CTV resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the CTV resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the Mt. Roose lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the University of California Riverside Genomics Institute, Core Instrumentation Facility. (Choi, Kingan).

Choi, then processed at UC Riverside (by Mananaker) using the HarvEST pipeline (http://harvest.ucr.edu) to remove contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank.

```

Qy      742  GCAAGTACCTTGCACCGGATTTTACGTGCAAAATTGATGAGCTTTGGAAGATCACT 801
Db      80  GCCAGTAGTCTCGCTCCAGATTGACGAGCAACATTTGACAGAGTGGAGAGTCAAC 139
Qy      802  CCGGTTATATGCTTGGAGTACTTGGCTTACCGCTTGATGATTAACGCTGCAAAAGA 861
Db      140  CCAGCTGTCTTGTGAACTTTTAAAGCTTTACCTTGTGTGTGATGATGATGATGATG 199
Qy      862  CTAAATGCTTTAAGCGGTGCGAAATATTTGTGTCTTGTGAGAGAGGTGAGATCA 921
Db      200  GAAGAAGAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
Qy      922  GCTCTTGTGTGGGCTTGAACCGCTGAGAACTTTATGATGATGATGATGATGATGATG 981
Db      260  GCAATGTGCTGGGAGATTTCACGTGATCTTTATGATGATGATGATGATGATGATGATG 319
Qy      982  GCTGTGAGAGGTTGATCTTTTGTGATGATGATGATGATGATGATGATGATGATGATG 1041
Db      320  TCACTGAGAGGTTTAACTTATTTCTGACACACCAATGATGATGATGATGATGATGAT 379
Qy      1042  GAAGTTACGAAGTTCACCTTGTGCTCAAGCTTTTATGTTAGAACCAACAC 1101
Db      380  GAAGCTTATGAGATGACCTTGTGATGATGATGATGATGATGATGATGATGATGATG 439
Qy      1102  CTTTACAGAGTCTGATTAAGCAATTTCAAGCACTTCAAGGCTTAAAGTATGCTATG 1161
Db      440  CTCAATGCAAGTCTGATTAAGCAATTTCAAGCACTTCAAGGCTTAAAGTATGCTATG 499
Qy      1162  GAGATTCGCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
Db      500  CGGACCTTGGCTCTATCTATATCTCTTGGAAAAATGATGATGATGATGATGATGATG 559
Qy      1222  AGGGAGCTGTGCACTGCTTATAGCAAAAGTATGATGATGATGATGATGATGATGATG 1281
Db      560  AGGGGCTATATGTTCACTGCTTGAAGTATGATGATGATGATGATGATGATGATGATG 619
Qy      1282  GACAGTGAAGATTCACAAATATGATGATGATGATGATGATGATGATGATGATGATG 1341
Db      620  GATAGTGAATAGTCCCAATACAAATACAGCAATGATGATGATGATGATGATGATGATG 679
Qy      1342  AATGATGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1401

```

Db 680 MAGGAGCTAATGACATATCTCTCGTCTGTAAAGCTTTAGACATGGCTGAGCA 739
 Qy 1402 GGGGTGCTCTTCTCTGCTTACAGACACCAAGATTTAAATCTCGGAGCTAC 1461
 Db 740 GAGGTGTTTCCCTGATTCAGGACACCAAGATTAACGTTCAACTGGTGTATAC 799
 Qy 1462 TATGATGATCTTAATGTTTGAAGTTACTTGAAAGAGTAGGTAGTTCAGGCTTCTCT 1521
 Db 800 TATGATGATCTTAATGTTTGAAGTTACTTGAAAGAGTAGGTAGTTCAGGCTTCTCT 859
 Qy 1522 TTAGCTGCTGCTGCAACTATGCGAAGATTGAGC 1556
 Db 860 TTAGCTGCGACGACCTGCAATGTTAGGCTTGAGC 894

RESULT 9
 B2437564/c 785 bp DNA linear GSS 13-DEC-2002
 LOCUS BONNR72TR_BO_1_6_2_KB.tot.Brassica oleracea genomic clone BONNR72,
 DEFINITION genomic survey sequence.
 ACCESSION B2437564
 VERSION B2437564.1 GI:26691135
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 785)
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
 Utterback,T.R., Mortman,J.R., White,O.R. and Town,C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its
 application to gene discovery and annotation in Arabidopsis
 Genome Res. 15 (4), 487-495 (2005)
 JOURNAL PUBMED 15805490
 COMMENT Other GSSs: BONNR72TF
 Contact: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..785
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BONNR72"
 /clone_1lb="BO_1.6_2_KB.tot"
 /note="Vector: pHD01, Site 1: BstXI, 1.6-2 kb sheared
 total DNA inserted into pHD01 using BstXI linkers"

ORIGIN
 Query Match 15.0%; Score 360.8; DB 9; Length 785;
 Best Local Similarity 86.4%; Pred. No. 2,1e-90;
 Matches 425; Conservative 0; Mismatches 57; Indels 10; Gaps 2;

Qy 505 AAGGTTCTGGGGCTCTCTGTATTTGCAAGAGGTGTGAGACTGAGATGTTCTTCGG 564
 Db 693 AAGGTTCTGGGGCTCTCTGTATTTGCAAGAGGTGTGAGACTGAGATGTTCTTCGT 634
 Qy 565 GTTGGAGAGCTCTGCTTAAGAGAGGTGCTTAAGTCTTAAAGAGATGTTGTTTA 624
 Db 633 GTTGGAGAGCTCTGCTTAAGAGAGGTGCTTAAAGTCTTAAAGAGATGTTGTTT 574
 Qy 625 GTTATGCGCTTGGCTTCTGATGTCTGAGGAGATCTATGAGATGATCCACTGAT 684
 Db 573 GTTATGCTGCTTGGCTTCTGATGTCTGAGGAGATCTATGAGATGATCCACTGAT 514

Qy 685 TTTATTACTGTTATGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGCA 744
 Db 513 TTTATTACTGAGATATGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGCA 454
 Qy 745 AGTAGCTTGCACCGGATTTACGTGACAAATTTATGATGACTTTGGAGAGATTCCTCG 804
 Db 453 ACCAGCTTGCACCGGATTTACGTGACAAATTTATGATGACTTTGGAGAGATTCCTCG 394
 Qy 805 C-GTTATGCTGAGAGCTACTTGGCTTACCGCTTGTGATGATTAAGCTGCGAAAAAGCT 863
 Db 393 CGATTATGATTTGAGACTCTTGGCTTACCTCTGGGGATGAT-----AAAGACA 343
 Qy 864 AATGCTTTAGCGGCTGTCGGAATATTTGTGCTGTTGAGAGAGTGAAGCATCAGC 923
 Db 342 AGATGCTTTAGCGGCTGTCGGAATATTTGTGCTGTTGAGAGAGTGAAGCATCAGC 283
 Qy 924 TCTTGTGGGGGTTTGAACCCGTGAGAAATTTATGATGAGGGCTTTTACGAATGACAC 983
 Db 282 TGTGTGTTGGTGTCTTACACGTGAGAAATTTATGATGAGGGCTTTTACGAATGACAC 223
 Qy 984 TGCTGAGCAGGT 995
 Db 222 ATCTGAGCAAGT 211

RESULT 10
 CX195205 624 bp mRNA linear EST 28-DEC-2004
 LOCUS 13-ED19365-021-009-J03-T7R ADIS-MP1Z 021 Brassica napus cDNA clone
 DEFINITION MP1Z1022J039Q, mRNA sequence.
 ACCESSION CX195205
 VERSION CX195205.1 GI:56842629
 KEYWORDS EST.
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 624)
 Jakoby,M., Lehmann,D. and Weishaar,B.
 direct submission to GenBank (ADIS-MP1Z 021)
 Unpublished (2004)
 Contact: Bernd Weishaar
 Chair of Genome Research
 Bielefeld University, Institute for Genome Research
 Universitaetsstrasse 25, D-33594 Bielefeld, Germany
 Tel: +49-521-106 6873
 Fax: +49-521-106 6423
 Email: bernd.weishaar@uni-bielefeld.de
 Insert Length: 624 Std Error: 0.00
 Plate: 9 row: J column: 3
 Seq primer: T7R CTATATCGACTCATATAGGA.
 Location/Qualifiers
 1..624
 /organism="Brassica napus"
 /mol_type="mRNA"
 /culivar="Express 617"
 /db_xref="GABI:1111775"
 /db_xref="taxon:3708"
 /clone="MP1Z1022J039Q"
 /issue_type="flowers and flower buds"
 /dev stage="young flowers and flower buds"
 /clone_1lb="ADIS-MP1Z 021"
 /note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; cDNA
 library from Brassica napus, strain Express 617; RNA was
 from young flowers and flower buds of two greenhouse
 plants; library was made at the Max-Planck-Institute for
 Plant Breeding Research, Cologne, Germany; cloning sites
 SalI-NotI, primer sites and orientation:
 T7-SalI-CCACCGTCCG-58rr-cDNA-polyA-CC-NotI-Sp6; made by
 Marc Jakoby 09/2000; PI: Bernd Weishaar. Sequence
 submission managed by RZPD/GABI-Primary database:

ORIGIN

http://gabi.rzpd.de"

Query Match 14.9%; Score 359.2; DB 8; Length 624;

Best Local Similarity 76.9%; Pred. No. 5.6e-90; Mismatches 93; Indels 63; Gaps 4;

Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;

1303 AGGAATCCAGCTATGTTGAGTTGTTTGGAGAAATCAATCGTATGCAATGATGAT 1362

2 AGAATCCAGCTATGTTGAGTTGTTTGGAGAAATCAATCGTATGCAATGATGAT 61

1363 CTCCTGAGCTATGCAAAATGTTGAGAACTGTGTCAGAGGTTGTTCTTCTAGGTT 1422

62 CTCCTGAGCTATGCAAAATGTTGAGAACTGTGTCAGAGGTTGTTCTTCTAGGTT 121

1423 AGAGACCCAAAGTAAAAATTTAACTGGGAGCTACTATGATGATGATGATGAT 1482

122 AGAGATACCAAAATTTAACTGGGAGCTACTATGATGATGATGATGATGATGAT 181

1483 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1542

182 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 241

1543 GCAAGATGGAAGCCGAGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1602

242 GCAAGATGGAAGCCGAGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 301

1603 CCTTCCGCTATGCAAAATGTTGAGAACTGTGTCAGAGGTTGTTCTTCTAGGTT 1662

302 CCTTCCGCTATGCAAAATGTTGAGAACTGTGTCAGAGGTTGTTCTTCTAGGTT 337

1663 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1719

338 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397

1720 GCAGAGCTGTAAGACCTCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1779

398 GCAGAGCTGTAAGACCTCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 431

1780 GTCTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1839

432 ----CAAGTCTAGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 487

1840 GAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1899

488 GAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 547

1900 CAGAGATGTTTCTTAAAA-----GAGCTCATCTTTTCAAGCAAGATGATGAT 1953

548 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607

1954 TCTATGAAATCTGA 1967

608 TCTACGCACTCTGA 621

RESULT 11

LOCUS CO075595 769 bp mRNA linear EST 15-JUN-2004

DEFINITION GR_Ba36G17.r GR_Ba Gossypium raimondii cDNA clone GR_Ba36G17, mRNA sequence.

ACCESSION CO075595

VERSION CO075595.1 GI:48745076

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 769)

AUTHORS Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wang, R. A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: rwing@genome.arizona.edu

Plate: 36 row: G column: 17.

FEATURES

source

1..769

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ba36G17"

/issue_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH10B"

/clone_lib="GR_Ba"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Clones plated/picked by Agt. More glycerol clones held in -80."

ORIGIN

Query Match 14.3%; Score 344.2; DB 7; Length 769;

Best Local Similarity 68.9%; Pred. No. 1.1e-85; Mismatches 213; Indels 9; Gaps 1;

Matches 491; Conservative 0; Mismatches 213; Indels 9; Gaps 1;

1 ATGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60

54 ATGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 113

61 CCGGCGACGACAAAGCTCCGAGCTGACCAACAGCTTAACTATGCTCCGACG 120

114 CACCTCTCTTAAACCTTCAGACCTTCAGACCTTCAGACCTTCAGACCTTCAG 173

121 AATGAGGCGACCGCTCTCTCTCCGACTTCAATT-----CACCTCCGATTCCT 171

174 AATGAGGCGACCGCTCTCTCTCCGACTTCAATT-----CACCTCCGATTCCT 223

172 TCCTCCCTCTGCGACCGACCAACACCGACCTCTCTCTGCGACCATCTATT 231

234 TCCT 233

232 GATGTCGCGAAGCGACCGACCTCCCATCCCATGATTTCTACAGGATTTAGAGCTCA 291

294 TCCT 353

292 ACACATTTCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 351

354 ACTCATTTCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 413

352 CAATTGCTTTGAGGACGCGCTTAAATCAAGCGGAGACAGATTCCTTCAAGCTGTC 411

414 CAATTGCTTTGAGGACGCGCTTAAATCAAGCGGAGACAGATTCCTTCAAGCTGTC 473

412 GAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471

474 GAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 533

472 GAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531

534 CGTGACATCATCTATCTCAAGTCCCTGCGACAAAGGTTCTTGACATTTGCGTGG 593

532 CAAGAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 591

594 CAAGAGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 653

592 TTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 651

654 CTGCCCAAGCAATTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 713

QY 652 TCGAGGAGCTGATGCTGATGATCCAGTATTTATCTGATTAGAGTT 704
 |||||
 DB 714 TCGAGGAGCTGATGCTGATGATCCAGTATTTATCTGATTAGAGTT 766

RESULT 12
 CX543653
 LOCUS
 DEFINITION 741 bp mRNA linear EST 12-JAN-2005
 UCRPT01_5_008 C05_T3 Poncirus trifoliata CTV-challenged cDNA
 library - UCRPT01-UCR2 Poncirus trifoliata cDNA clone

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

EST
 CX543653.1 GI:57570678

Poncirus trifoliata
 Poncirus trifoliata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eustosida II; Sapindales; Rutaceae; Poncirus.

1 (bases 1 to 741)
 Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
 Manamaker, S., Choi, Y. and Kiang, T.

Development of EST Resources and New Genetic Markers for California
 Citrus - Poncirus trifoliata CTV-challenged phloem - UCRPT01-UCR2
 Unpublished (2004)

Contact: Mikeal Roose
 Department of Botany & Plant Sciences, University of California
 Riverside, CA, 92521-0124, USA

Tel: 9097874137
 Fax: 9097874437

Email: mikeal.roose@ucr.edu
 Seq primer: T3.

Location/Qualifiers

1..741
 /organism="Poncirus trifoliata"
 /mol_type="mRNA"
 /cultivar="Pomeroy OP"
 /db_xref="taxon:37690"
 /clone="UCRPT01_008_T3_C05"
 /isue_type="bark (with phloem)"
 /dev_stage="10 - 30 cm shoots"
 /lab_host="E. coli TUC121"
 /clone_1lb="Poncirus trifoliata CTV-challenged cDNA
 library - UCRPT01-UCR2"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
 greenhouse at University of California, Riverside. The
 action was an open-pollinated (very probably selfed)
 seedling of Poncirus trifoliata cv Pomeroy that was
 selected as homozygous for the CTV resistance gene. The
 rootstock was sweet orange infected with citrus tristeza
 virus (CTV) isolate T514 over 1 year before sampling (CTV
 infects sweet orange, but not genotypes carrying the CTV
 resistance gene. Shoots 10-30 cm long were harvested in
 October 2000, and the green phloem (bark) was removed and
 frozen quickly in dry ice. Total RNA was extracted using
 Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
 library was made, and 0.5 million primary lambda cDNA
 clones were in vivo excised to give a population of
 phagescript SK(-) phagemid. All steps to this point were
 performed in the ML Roose lab at the University of
 California, Riverside by X. Ye. Phagemids were plated,
 plasmid DNA purified, cDNA clones archived, and DNA
 sequences determined bi-directionally using an ABI3730 at
 the University of California Riverside Institute of
 Integrative Genome Biology Genomics Core Instrumentation
 Facility, (Choi, Kiang). Chromatogram files were
 downloaded by FTP by Close, then processed by Manamaker
 (Close lab) using the HarVest pipeline
 (http://harvest.ucr.edu) to remove vector and cloning
 oligo sequences and various contaminants, and to trim to a
 high quality region. Sequences that retained a phred 17

region of at least 100 bases were assembled, then chimeras
 were removed following manual inspection of assemblies
 (Close, Roose, Manamaker). Sequences that survived all
 removal steps were submitted to GenBank."

ORIGIN

Query Match 13.9%; Score 335.4; DB 8; Length 741;
 Best Local Similarity 73.3%; Pred. No. 3.4e-83;
 Matches 445; Conservative 0; Mismatches 156; Indels 6; Gaps 1;

QY 98 CTACACATATTCGCTCCGACGAAATGGGCGGACCGCTTCCTCCGACTCAATTTC 157
 |||||
 DB 138 CCACCATCTCTCTCCGACGAAATGGGCGGACCGCTTCCTCCGACTCAATTTC 197
 158 CCTCCGATTCCTC 217
 198 CCACCGCGACCAATTCCTC 257
 QY 218 TGCACCATCTATTTGATGCTCCGAAAGCGCAGCTCCCATCCCATTTGATTCTACAG 277
 258 CCCTCC-----TACTCCACCAACCGCACGATTCATCCCATCGACTTCTATCAAG 311
 DB 278 TATTAGAGACTCAACACATTTCTTACCGATGGAATCAAGAGAGATTCGAGGAG 337
 312 CGTTGGAGCGGAGACTCAATTTCTTGGAGATGGAATTAAGAGGCTTATGAACTAGGA 371
 QY 338 TTTCGAAACCGCGCAATTCGTTTCAGGACGACGCTTTAATCAAGCCGAGACAGATTC 397
 372 TTTCCTAACCACTCTCACTGACGGGTTTACGCTGATGATGACCTTACGAGATTC 431
 DB 432 TTCAAGCTCTCTGTGAACCTTAGCTAATGCTAGCTCTAAGAGAAATCAATCAAGGA 491
 QY 458 TTCTTATGATGAAGAAGCTACAGTATCACTGATGTTCTTGGATTAAGTTCCGAGG 517
 492 TCGCTATATACCAATGCGAATACATCTCACTGAATTTCTTGGCAAGGTTCTGAG 551
 QY 518 CTCTCTGTGATTCGAAGAAGGTGAGACTGAGATGATTTCTTGGGTTGGTGAAGCTC 577
 552 CATTGCTTGTGTCGAAGAGCTGGGAGACGAGAGTATGCTGAGAAATTTGAGAGAAATC 611
 DB 578 TGGTAAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTTTATGATGAGCGCTTG 637
 612 TGTGAGAGAGAGAGACTTCCCAAGTCTTTCAAGAAAGATGTTGCTTGGGAGTGGCACTTG 671
 QY 638 CGTTTCTGATGTTCTGAGGAGATGCTATGCGATTCAGCTGATTTTATTAATCTGTT 697
 672 CCTATGTTGACATTTCCAGGAGATGCTATGAGCAATTAATCCGCTGATTAATGAGGCT 731
 DB 698 ATGAGATT 704
 DB 732 GTGAGAT 738

RESULT 13

BU046755 631 bp mRNA linear EST 26-AUG-2002
 LOCUS
 DEFINITION PP_LBA0027104F Peach developing fruit mesocarp Prunus persica cDNA
 clone PP_LBA0027104F, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

EST.
 Prunus persica (peach)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eustosida II; Rosales; Rosaceae; Amygdaloideae; Prunus.

1 (bases 1 to 631)
 Callahan, A., Palmer, M., Main, D., Wing, R. and Abbott, A.
 Peach Model Genome for Rosaceae
 Unpublished (2002)
 Contact: Abbott, A.

Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu

Total High Quality bases = 523
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

Source

1. 631

/organism="Prunus persica"

/mol_type="RNA"

/culivar="Loring"

/db_xref="taxon:3750"

/clone="PP_LBA0027104f"

/tissue_type="Mesocarp"

/lab_host="E. coli"

/clone_lib="Peach developing fruit mesocarp"

/note="Vector: Bluescript II SK(-); Site 1: EcoRI;

Site 2: XhoI; authority=Prunus persica L. Batsch; The

sequence has been trimmed to remove vector sequence and

contains a minimum of 100 bases of phred value 20 or

above. For more details on library preparation and

sequence analysis go to

http://www.genome.clemson.edu/projects/peach. To order

this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match

Best Local Similarity 71.5%; Pred. No. 2,4e-81;

Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;

13.7%; Score 328.8; DB 5; Length 631;

1

913 GAGCATCAGCTCTTGGGGGTTTGAACCGTGAGAGTTTATGAGGCTTTTGA 972

1 GCAGTTGCAATTCGCGGGGNGATTCACTACGTGAAATTTCAATGAAAGGCTTCTTG 60

973 CGAATGACGCTGTGAGAGGTTGATCTTTTGTACTACCCCAAGCAATATTCGCA 1032

61 CATATGACTGACGAGGTTGATTTATTTGTACTACCCCAAGCAATATTCGCGCA 120

1033 GAGTCATTTGAGTTTACGAAGTTGCACTTCTTGTGCTCAAGCTTTTATTTGTA 1092

121 GAAAGCTTTGAGTTTATGAGGTTGCTTTCGCTTCTTCTCAAGCTTTTGTGTA 180

1093 AAGCCACACTTTTACAGATGCTGATAGCAATTCAGCAACTTCAGAGCTTAA 1152

181 AAACCTCATCACTTCAAGATGCTGAAACCTTTCAGAACTTCAGAGCTTAA 240

1153 ATGGCTATGAGATTCCTGAGATTTGTATGATACAGGAATTAATGGAGATTA 1212

241 ACAGCTGTAGACATTCCTTGAACAATATTAACAAAGAGCAATGATGATG 300

1213 GGTCTGAAAGGAGGACTCTGTGCACTTATAGGCAAAAGTTGATGATCCGAT 1272

301 GCTTTGAGAGGAGGAGGCTTTCATCTTCTTAGGGGACCTTGATGACATTC 360

1273 TTGGGCTTACAGATGAGATTCACATATAGAAATCAGCTATTTGAGTTT 1332

361 TTGGGCTTACAGATGAGATTCACATATAGAAATCAGCTATTTGAGTTT 420

1333 GAGAAATTCAGATGAGTATG-----ACAAATGATGATCTCCCTGAGATTA 1380

421 GAGAAATTCAGATGAGTATG-----ACAAATGATGATCTCCCTGAGATTA 1440

1381 TTGTTGAAACCTGTTGAGAGGTTGCTTTCTTCTGAGTTTACAGAGCAATTA 1440

481 CTATTTGAGAGCTGTTGAGAGGTTGCTTTCTTCTGAGTTTACAGAGCAATTA 540

1441 AAATTTAACTCGGGAGCTATGATGATCTTATGTTTGAATTACTTGAAGAG 1500

541 GAGTTGAGAGCTGAGAGCTATGATGATCTTATGATGATCTTGAAGAGAGCT 600

QY 1501 GAGTAGTTCAGGAGTTCTCTTACGTCTG 1531
DB 601 GATGCACTAATGTTTACCTTACGTCTG 631

RESULT 14

CN918842

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

SEQUENCING FACILITY

The Horticulture and Food Research Institute of New Zealand Ltd

120 Mt Albert Rd, Mt Albert, Auckland, New Zealand

Tel: 00 64 09 815 4200

Fax: 00 64 09 815 4201

Email: est@hortresearch.co.nz.

Location/Qualifiers

1. 703

/organism="Malus x domestica"

/mol_type="RNA"

/db_xref="taxon:3750"

/clone="ABP006290"

/tissue_type="Root tips (distal 1.5 cm)"

/note="Vector: Bluescript SK(-); Library sequenced by

Genes Research & Development"

ORIGIN

Query Match

Best Local Similarity 71.5%; Pred. No. 1e-80;

Matches 450; Conservative 0; Mismatches 164; Indels 15; Gaps 1;

13.6%; Score 326.6; DB 7; Length 703;

1

74 AGCTCGACGCTAGCAACCACTTCACTATCTGCTCGCAGCAATGAGCGGAC 133

75 ACCACGGAAACCCCAAAAGCTCCCTCCCACTGCTTCCGACAAATGGCGAGC 134

134 GTCTTCTCGGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 181

135 GCTCTCTCGGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCT 194

182 ---TCGCAACGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 228

195 CCTCTCTCTCGGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 254

239 CCGAAGCGCAGCTCCCATCCCATTTGATTTCAAGAGTATTAAGAGCTCAACAT 298

255 CCGAGCGCAGCTCCCATCCCATTTGATTTCAAGAGTATTAAGAGCTCAACAT 314

299 TCTTAACCGATGAAATCAAGAGCTTGAAGCTGAGTTTGAAGAGCTCAACAT 358

315 TCTCTCGGAGCGGATTAAGAGGCGGATTAAGAGCTGAGTTTGAAGAGCTCA 374

359 GTTTAGCGAGCGGCTTTATCAAGCGGAGAGAGATTTCTTAAGCTGTTGAAAG 418

375 GCTTACCGAGAGGCGGCTTTATCAAGCGGAGAGAGATTTCTTAAGCTGTTG 434

419 TGTCTATCTCGGCTTGAAGAGAGTCAATGAAGCTTCTTGAAGAGAGAGCT 478

435 TAGCCGACCGCGCTCCGAGAGAGAGTCAACAGAGAGCTTTCGAGAGAGAGAG 494

QY 479 AAGTATCATCGAATGTTCTCTGGGTAAGATTCCTGGGGGCTCTGATATTTGAAGAAG 538

Db 495 CCAATATCACTCAAGATCTCTGGGAAAGATTCCTGAAGCTCTGTGCTGCTGCAAGAG 554

QY 539 GTGTGAGACTGAGATAGTCTTCTCGGGTGTGTGAGGCTCTGTGTTAAGAGAAGTTGCTTA 598

Db 555 CTGGGAAGACTGAGCTGGTCTTTCAAATTTGGGGAAGAATTGCTAAGAGAAGGCTGCCCA 614

QY 599 AGTCGTTTAAAGCAAGATGTGGTTTAAGTATTAAGGCGCTTGCCTTCTCATGTCTGAGGG 658

Db 615 AGTCCTTTAAGCAAGATGTGTTTGTGCATGAGCACTTGCTTAATGTGACATGTGAGGG 674

QY 659 ATGCTATGGCAATGGATCCACCTGATTTT 687

Db 675 ATGCATGAAATGTCCCGCGCTGATTTT 703

RESULT 15	415 bp	mRNA	linear	EST 10-FEB-2005
BP785511/c				
LOCUS	BP785511			
DEFINITION	BP785511 RAFL7 Arabidopsis thaliana cDNA clone RAFL07-95-D11 3',			
	mRNA sequence.			

ACCESSION	BP785511	
VERSION	BP785511.1	GI:59261187
KEYWORDS	EST.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	

REFERENCE
AUTHORS

1 (bases 1 to 415)
Seki, M., Ihida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Iida, K., Enju, A., Sakurai, T., Aikawa, T., Carrinno, P., Fukuda, S., Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and Shinozaki, K.

TITLE Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
JOURNAL Unpublished (2005)
COMMENT Contact: Motoaki Seki

Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998, 2002). This clone is in a modified pBluescript vector.
Please visit our web site (<http://pfweb.gsc.riken.jp> and <http://large.gsc.riken.jp>) for further details.
reversed clone.

FEATURES	Location/Qualifiers
source	1. .415

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/cclone="RAFL07-95-D11"
/dev_host="roseette plants"
/lab_host="DH10B"
/cclone_lib="RAFL7"
/note="Site 1: BamHI; Site_2: SalI; subjected to cold-treated (1, 2, 5, 10, 24 hr)"

```

ORIGIN				
Query Match	13.3%	Score 321,	DB 3;	Length 415;
Best Local Similarity	100.0%	Pred. No. 3.4e-79;		
Matches 321;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy 2086 CCTGATCACCAGATGAAATGTTTACCAGAGCTTTTGGATGGCGCAATGCTGAAGTTGG 2145

Db 415 CCTGATCACCAGATGAAATGTTTACCAGAGCTTTTGGATGGCGCAATGCTGAAGTTGG 356

Qy 2146 ACTGACAGAGCAGCTGAACTGCCGACCTTGGGTGTTTATGATTTATACCTGTTGAAA 2205

Accession	Sequence	Position
Db	355 ACTGACAGAGCAGCTGAAATCGCCGAGCTTGCGTTGTTATGATTATATACACTGTGGAAA	296
QY	2206 CTATCTGTGGACAGTGTGACAGCTCTCAGCAGATGGAACCCGATGCTCTGGTGGAAAGCACT	2265
Db	295 CTATCTGTGGACAGTGTGACAGCTCTCAGCAGATGGAACCCGATGCTCTGGTGGAAAGCACT	236
QY	2266 CTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCTCATCCAGAAAACAAATGCTACTGATGTC	2322
Db	235 CTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCTCATCCAGAAAACAAATGCTACTGATGTC	176
QY	2326 AGAACCTTACACACAGATACGAAATTTCTGGTCCAAAGTACAGGCTGGAAAATCACTGAA	2385
Db	175 AGAACCTTACACACAGATACGAAATTTCTGGTCCAAAGTACAGGCTGGAAAATCACTGAA	116
QY	2386 GGCCTCTGTTCTTGCATCATTA	2406
Db	115 GGCCTCTGTTCTTGCATCATTA	95

Search completed: February 21, 2006, 15:40:23
Job time : 6314 secs

QY		2386	GGCTGTGTTCTTGCAATCATAA	2406
Db		115	GGCTGTGTTCTTGCAATCATAA	95